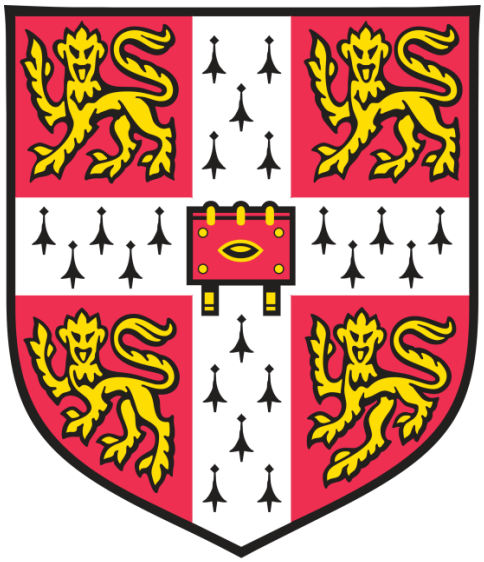
Selective Protein Functionalisation at Methionine

Jennifer E. Nelson

Murray Edwards College

University of Cambridge



This dissertation is submitted for the degree of Doctor of Philosophy

Department of Chemistry

University of Cambridge

Lensfield Road

Cambridge

CB2 1EW

April 2019

Declaration

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy. It describes the work carried out in the Department of Chemistry from October 2015 to March 2019. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

Jennifer Nelson

11th April 2019

Statement of Length

This dissertation does not exceed the word limit of 60 000 as set by the Degree Committee for the faculty of Physics and Chemistry.

Jennifer Nelson

11th April 2019

Acknowledgements

Firstly, I would like to thank Professor Matthew Gaunt; I feel that I have developed personally and achieved so much more than I could have imagined during my PhD and that is due to the opportunities and support I have received in this group. I am so grateful to have worked on such an interesting project and for the platform this experience has given me for my future career.

Thank you also to Mike, who I had the privilege of working with during the first two years of my PhD. I appreciate your patience and guidance, you taught me so much, and it was a pleasure to work together.

To Patrick W., I doubt that I would have survived my first year were it not for your friendship. You have been an important part of my time in Cambridge and I thank you for that. I'd like to thank Patrick D., Antonio and Henry for the craic, for bringing so much fun to the lab and for always being keen for boffice. Thank you also to the rest of the Gaunt group, I have been so lucky to have been able to work with some of my best friends every day, and it has made this experience so much more enjoyable.

To Patrick D. and Patrick W., thank you for your tireless efforts in proof reading this thesis, I am extremely grateful. Thank you also to Henry and Antonio for taking the time to proof read this work.

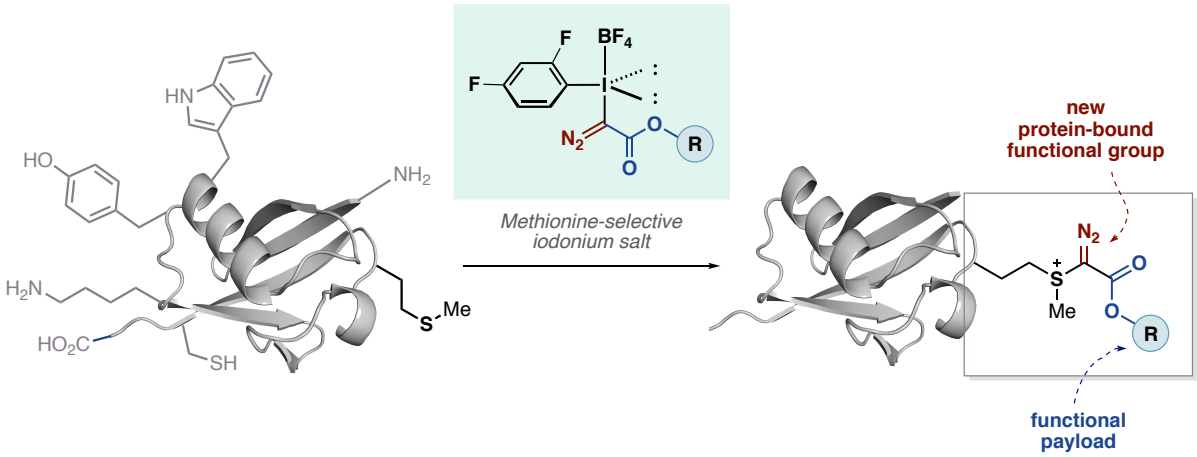
I would like to acknowledge the technical staff, Nic, Naomi and Carlos for their support, the NMR service, Duncan, Andrew and Peter, and the mass spectrometry service, Dijana and Asha, for their help and advice. Thank you also to AstraZeneca and the EPSRC for their generous financial support. I am grateful to Professor Jason Chin, Dr Nicolas Huguenin and Dr Mark Skehel (MRC Laboratory of Molecular Biology) for their assistance with FPLC purification and Dr Hilary Lewis and Dr Matthew Edgeworth (AstraZeneca/MedImmune) for carrying out MS/MS analysis.

Finally, thank you to my family, and to Henry, for their unwavering love and support, and their continued confidence in me over the last three and a half years.

Abstract

The development of novel bioconjugation strategies for the functionalisation of polypeptides and proteins has been of great benefit to the fields of chemical biology and medicine. Selective protein functionalisation has enabled the development of technologies such as targeted drug delivery, enzyme activity profiling and imaging of cells *in vivo*. Many strategies for bioconjugation target the nucleophilic side chains of the amino acids cysteine and lysine. However, many strategies remain sequence specific and some require pre-installation of recognition units to achieve high selectivity.

Due to the low natural abundance and ancillary protein function of methionine, selective functionalisation of this amino acid presents an opportunity to develop a strategy complementary to current technologies. Strategies targeting methionine remain under-explored, with few truly biocompatible methods reported to date. This thesis describes the development of a novel bioconjugation method employing hypervalent iodine reagents for the selective functionalisation of methionine residues. The broad scope and excellent biocompatibility of the reaction is demonstrated, with a number of different polypeptide and protein substrates tolerated. Additionally, through the synthesis of different hypervalent iodine reagents, varying the ester group of the iodonium salt (R), a number of different payloads can be transferred to polypeptides and proteins.



The highly reactive diazo group which is introduced using this methodology can be exploited in subsequent bioorthogonal transformations. Firstly, phosphine reagents can be employed for the cleavage of the methionine label, in a stimuli-responsive reversal of the conjugation. Secondly, a visible light-mediated reduction was developed, forming trialkylsulfonium

products which exhibited improved stability. Finally, a *C*-benzylation has been developed using photoredox catalysis to enable functionalisation of this diazo moiety.

The methionine bioconjugation strategy has also been used in tandem with a literature procedure for the functionalisation of tryptophan residues. Through the combination of two bioconjugation techniques, a simultaneous dual functionalisation of polypeptides has been developed, forming dually- or triply-functionalised scaffolds in a single step from native polypeptides.

Abbreviations

Å	angstrom
Ac	acetyl
ADC	antibody-drug conjugate
Ala	alanine
aq.	aqueous
Ar	aryl
Asp	aspartic acid
Bn	benzyl
Boc	*tert*-butyloxycarbonyl
br.	broad
Bu	butyl
bpy	2,2'-bipyridine
Bz	benzoyl
C	celsius
cat.	catalyst
Cbz	carboxylbenzyl
CD	circular dichroism
cm^{-1}	wavenumbers
CuAAC	copper(I)-azide-alkyne [3+2] cycloaddition
Cy	cyclohexyl
Cys	cysteine
Da	daltons
DARPin	designed ankyrin repeat protein
DBCO	dibenzocyclooctyne
decomp.	decomposition
Dha	dehydroalanine
DIPEA	*N,N*-diisopropylethylamine
DMF	*N,N*-dimethylformamide
DMP	Dess-Martin periodinane
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dtbbpy	4,4'-di-*tert*-butyl-2,2'-bipyridine

e	electron
E	electrophile
Et	ethyl
eq.	equivalent
ESI	electrospray ionisation
fac	facial
Fmoc	fluorenylmethyloxycarbonyl
FPLC	fast protein liquid chromatography
FRET	Forster resonance energy transfer
g	gram
GFP	green fluorescent protein
Gly	glycine
HAT	hydrogen atom transfer
HCTU	O-(1H-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
His	histidine
hr	hour
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
Hz	hertz
Int	intermediate
i	iso
IR	infrared
J	coupling constant
k	rate constant
k	kilo
ketoABNO	9-azabicyclo[3,3,1]nonan-3-one-9-oxyl
L	litre or neutral ligand
LCMS	liquid chromatography mass spectrometry
LED	light emitting diode
Leu	leucine
Lys	lysine
μ	micro
M	metal or molar
m	milli

Me	methyl
Mes	mesityl
Met	methionine
MHz	megahertz
min	minute
mol	mole
m.p.	melting point
MS	mass spectrometry
Ms	methanesulfonyl
m/z	mass to charge ratio
NHS	*N*-hydroxysuccinimide
NMR	nuclear magnetic resonance
Nu	nucleophile
PEG	polyethylene glycol
PET	positron emission tomography
Ph	phenyl
Phe	phenylalanine
PMDETA	*N,N,N',N'',N''*-pentamethyldiethylenetriamine
ppm	parts per million
ppy	2-phenylpyridinato
Pr	propyl
Pro	proline
quant.	quantitative
R	undefined group
R_f	retention factor
RNA	ribonucleic acid
ROS	reactive oxygen species
rt	room temperature
s	seconds
SAM	*S*-adenosyl methionine
sat.	saturated
Ser	serine
SM	starting material
SPAAC	strain-promoted azide-alkyne [3+2]cycloaddition
t	tertiary

TAMRA	5-carboxytetramethylrhodamine
TCEP	tris[2-carboxyethyl]phosphine
TEMPO	(2,2,6,6-tetramethylpiperidin-1-yl)oxyl
Tf	trifluoromethanesulfonate
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIC	total ion count
TMS	trimethylsilyl
Tris	trisaminomethane
Trp	tryptophan
Ts	tosyl
Tyr	tyrosine
UV	ultra-violet
V	volts
ν_{max}	infrared absorption
W	watt
X	halogen or anionic ligand

Table of Contents

Declaration i

Statement of Length i

Acknowledgements ii

Abstract iii

Abbreviations v

Table of Contents ix

1 Introduction 1

1.1 Introduction to protein modification 1

1.2 Traditional bioconjugation strategies 3

1.2.1 Established bioconjugation reactions of cysteine 3

1.2.2 Recent developments in cysteine functionalisation 6

1.2.3 Bioconjugation of lysine 14

1.3 Methionine bioconjugation 18

1.4 Conclusion and outlook 31

2 Development of a methionine-selective bioconjugation strategy 32

2.1 Hypervalent iodonium salts for bioconjugation 32

2.1.1 An α-aryliodonio diazo compound as a reagent for methionine bioconjugation 36

2.2 Previous work 37

2.3 Project aims 42

2.4 Results and discussion 44

2.4.1 Optimisation of methionine bioconjugation and isolation of sulfonium products 44

2.4.2 Structural analysis of β-sulfonium α-diazoester conjugates 47

2.4.3 Scope of iodonium salt transferring reagents 52

2.4.4 Stability studies of methionine sulfonium conjugates 54

2.4.5 Mass spectral characteristics of β-sulfonium α-diazoester-containing peptide and protein conjugates 58

2.4.6 Exploiting the reactivity of the diazo motif in β-sulfonium α-diazoester conjugates 61

2.4.7 Summary 89

3 Synthesis of multi-functionalised polypeptides and proteins 90

3.1 Introduction 90

3.1.1 Double labelling using identical amino acids with different reactivity 91

3.1.2 Double labelling using two different amino acids 94

3.2 Project aims 99

3.3 Results and discussion 100

3.3.1 Dual-labelling of methionine and tryptophan residues 100

3.4 Future work 120

3.5 Summary 122

4 Conclusions and Outlook 123

5 Experimental 129

5.1 General Information 129

5.2 Experimental procedures for the synthesis of hypervalent iodine reagents 132

5.3 Experimental procedures for polypeptide and protein functionalisation 145

5.3.1 Substrate scope for polypeptide and protein functionalisation 147

5.3.2 Substrate scope of hypervalent iodine reagents for polypeptide functionalisation 167

5.4 Cyanogen bromide mediated backbone cleavage of Aviptadil 175

5.5 Experimental procedures for exploitation of diazo motif in β-sulfonium α-diazoester conjugates 176

5.5.1 Experimental procedures for phosphine promoted cleavage of methionine label 176

5.5.2 Experimental procedures for photochemical reduction of α-sulfonium diazoesters 179

5.5.3 Experimental procedures for a photoredox-mediated bioorthogonal transformation at β-sulfonium α-diazoesters 184

5.6 Experimental procedures for the multi-functionalisation of polypeptides 199

5.6.1 Dual labelling of methionine and cysteine 199

5.6.2 Dual labelling of methionine and tryptophan 201

5.6.3 Substrate scope for the multi-functionalisation of polypeptides 207

6 References 214

Appendix I 230

Appendix II 256

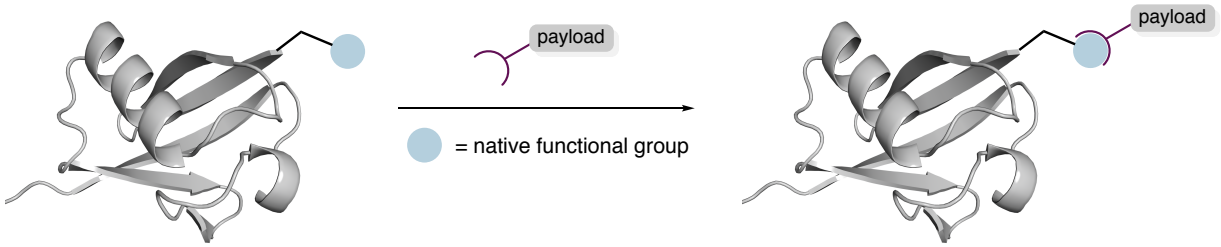
1 Introduction

1.1 Introduction to protein modification

Methods for the modification of peptides and proteins have proven integral to the advancement of chemical biology and medicine. Novel strategies to introduce exogenous functionality in a selective fashion have enabled the investigation of biomacromolecules and have provided insights into cellular mechanisms.[1,2] The resulting functionalised species can exhibit control over chemical and biological behaviour, providing a platform for the development of targeted therapeutics and cellular imaging agents.[3]

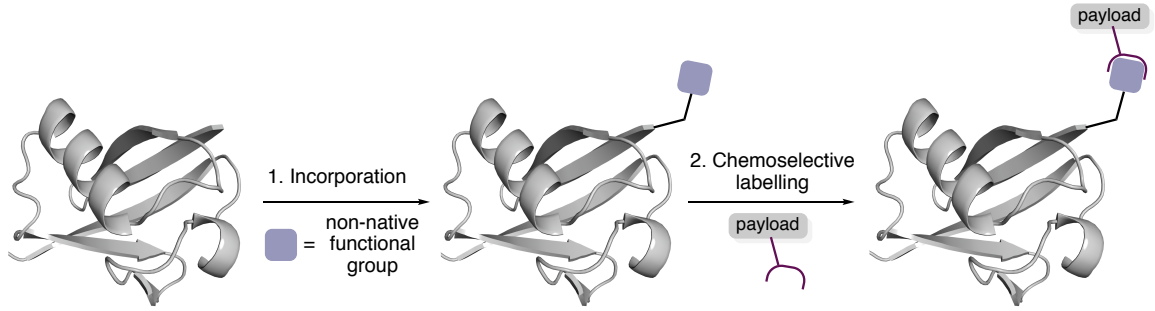
The development of selective chemical modification strategies for proteins must overcome challenges unique to biological systems, including the abundance of competing, reactive functionality capable of unwanted side reactions and the need for aqueous conditions.[4] In order for reactions to be considered biocompatible they must strive to fulfil several criteria. Namely, they should proceed with almost perfect conversion, yielding homogeneous, stable products under biologically ambient conditions (aqueous solvent, pH 6-8, < 37 °C).[4,5,6] Additionally, it is important that reactions should not perturb the protein architecture so as to preserve the integrity of the biomolecule's structure and function.[7] For reactions carried out *in vivo,* it is desirable for the kinetics to be comparable to that of an enzymatic timescale ($k = 10^4 – 10^5$ $M^{-1}s^{-1}$)[8] to ensure that the reactions can occur at low concentrations and can compete with, or be used to investigate, cellular processes. Many existing methodologies for the synthetic modification of small molecules prove incompatible with polypeptide and protein systems. Therefore, the development of new, biocompatible reactions are required.

Despite the stringent demands of biological systems, a large number of reactions that meet many or all of these criteria have been developed.[2] The native chemical functionality of peptides and proteins can be targeted using a set of techniques termed 'bioconjugation' (Scheme 1). These approaches typically exploit the reactivity of the side-chains of canonical, proteinogenic amino acid residues present on the surface of proteins.[4]



Scheme 1 Bioconjugation to introduce a functional payload to a protein

In contrast, bioorthogonal chemistry exploits the reactivity of non-native functionality previously introduced into a biomolecule. Defined as reactions in which both starting materials and products do not interfere with biological processes,[4] bioorthogonal transformations are generally achieved in two steps.[9] Incorporation of a non-native functional group is followed by treatment of the system with a probe molecule bearing functionality which is complementary to the introduced moiety. These two coupling partners can undergo a secondary, chemoselective reaction to form the desired product (Scheme 2).[4,9] Incorporation of the non-native group is most commonly achieved through genetic encoding[10,11] or metabolic labelling,[12] but may also be accomplished *via* an initial bioconjugation reaction at a native residue.



Scheme 2 Two step bioorthogonal functionalisation of a protein; 1. Incorporation of non-native functionality through metabolic labelling, genetic encoding or bioconjugation reaction, 2. Chemoselective labelling of non-native functional group

Progress in the field of protein modification has enabled both the *in vitro* and *in vivo* functionalisation of biomolecules.[2,6,13] These advances have greatly benefited chemical biology and medicine; enabling targeted drug delivery using antibody-drug conjugates,[14] enzyme activity profiling[15,16] and imaging of cells *in vivo*[17] (Figure 1). The aim of this introduction is to inform the reader of key approaches to bioconjugation, focusing on important seminal work and recent developments for the targeting of the most-commonly modified residues, demonstrating the need for novel and distinct bioconjugation strategies. For a more detailed insight into bioorthogonal transformations, the reader is directed to a number of reviews.[4,8,12,18,19]

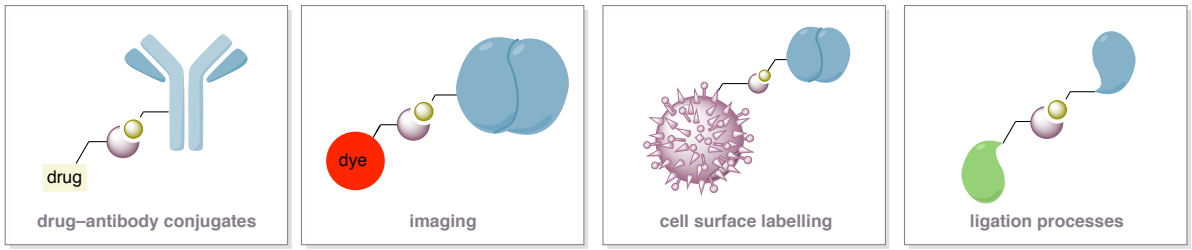
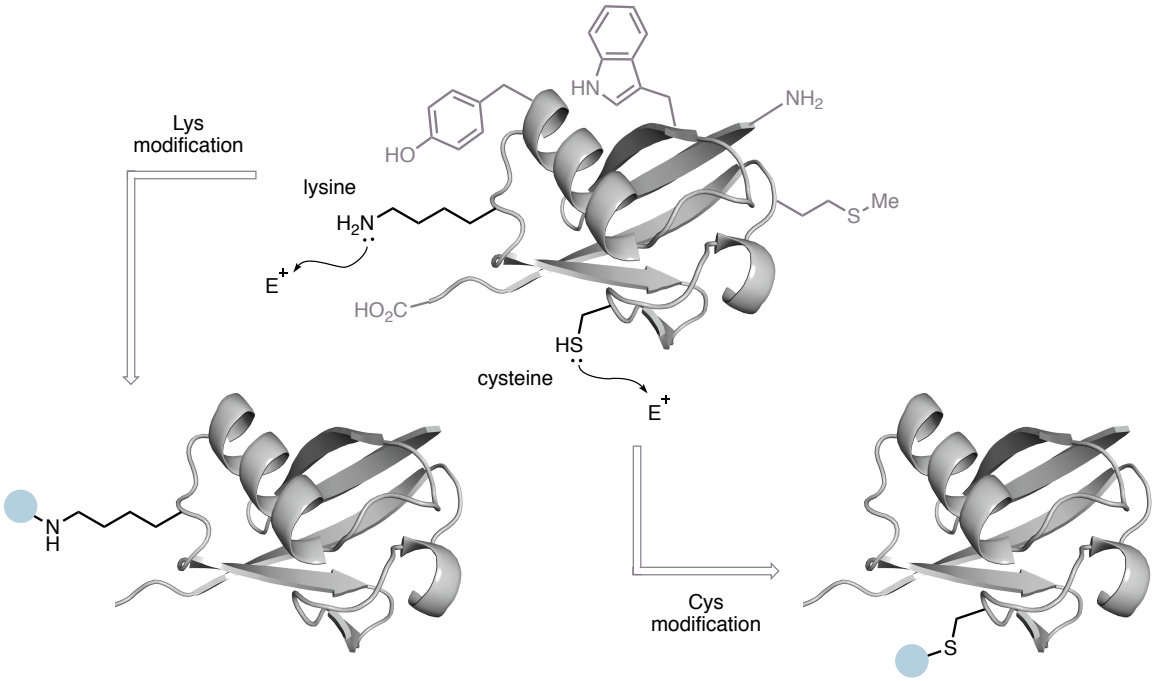


Figure 1 Potential biological applications of modified polypeptides and proteins

1.2 Traditional bioconjugation strategies

The most common bioconjugation methods target the ionisable and nucleophilic amino acids cysteine and lysine through reaction with a range of electrophiles (Scheme 3).[20]



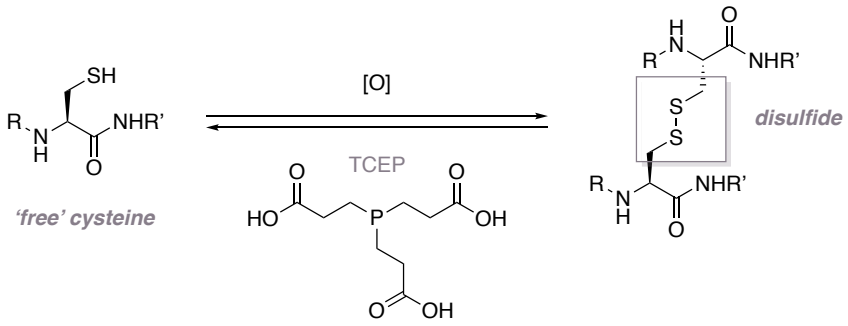
Scheme 3 Two of the most common strategies for bioconjugation: modification of cysteine and lysine through nucleophilic addition or substitution reaction with electrophiles

1.2.1 Established bioconjugation reactions of cysteine

Owing to the strongly nucleophilic nature of its sulfhydryl side chain,[21] cysteine is the most explored residue for bioconjugation. With initial reports as early as 1935,[22] strategies targeting cysteine are now routinely practised, with some widely acknowledged as the benchmark of protein bioconjugation. Possessing a pKa of approximately 8.5, at physiological pH the nucleophilic thiol of cysteine predominantly exists as the deprotonated thiolate form, which

surpasses the reactivity of any other nucleophilic function in proteins.[21] Bioconjugation is usually achieved through nucleophilic addition or substitution with the thiolate anion.[2] Furthermore, as a relatively rare amino acid (natural abundance ~3% in vertebrates),[23] cysteine can often be used to carry out single-site modifications. This is important as it facilitates the formation of homogeneous, well-defined products which are particularly vital for use as targeted therapeutics.[24]

Cysteine residues serve a unique and important structural role in proteins through the formation of disulfide bonds between the thiol groups of two side chains (Scheme 4).[25] The secondary and tertiary structure of many proteins can be influenced by these stabilising interactions.[26] As cysteine residues are generally present in the oxidised disulfide form, solvent-exposed free cysteine residues required for many bioconjugation strategies are rarely found in proteins.[27] This can be overcome through the generation of free cysteine *via* the reduction of disulfide bonds with reagents such as tris[2-carboxyethyl]phosphine (TCEP),[28,29] or through the incorporation of free cysteine *via* recombinant expression.[30] However, considering their important structural role, reduction and functionalisation of disulfide bonds can cause structural perturbation in proteins, which may prove detrimental to protein function.[31] Nevertheless, cysteine-selective modification strategies remain the most widely exploited in the field of bioconjugation.

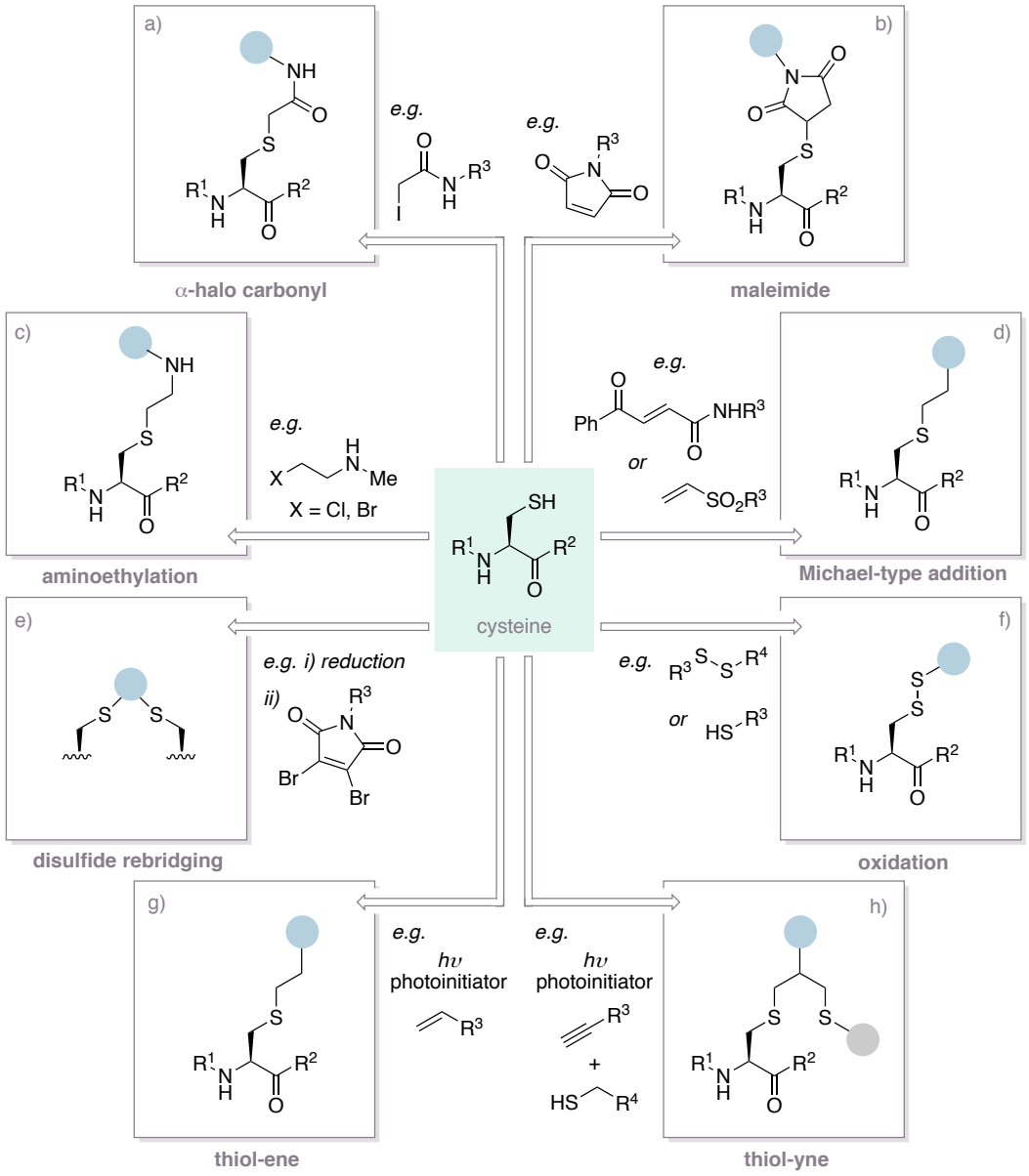


Scheme 4 Oxidation of cysteine within biological systems forms disulfide bonds which are pivotal to protein structure and function. Reduction with TCEP regenerates the free thiol group

Illustrated in Scheme 5 is a brief overview of established transformations targeting cysteine for polypeptide and protein functionalisation. The most widely used strategies involve Michael addition to maleimide reagents[32–36] due to the very fast kinetics (second order rate constant = 734 $M^{-1}s^{-1}$)[37] and high selectivity of these reactions (Scheme 5b).[3,5,18] The prevalence of this approach has led to the commercialisation of a number of maleimide-derived reagents, and some antibody-drug conjugates made through cysteine-maleimide conjugation are now approved drugs.[38] Other commonly employed reactions include the use of α-halo carbonyl

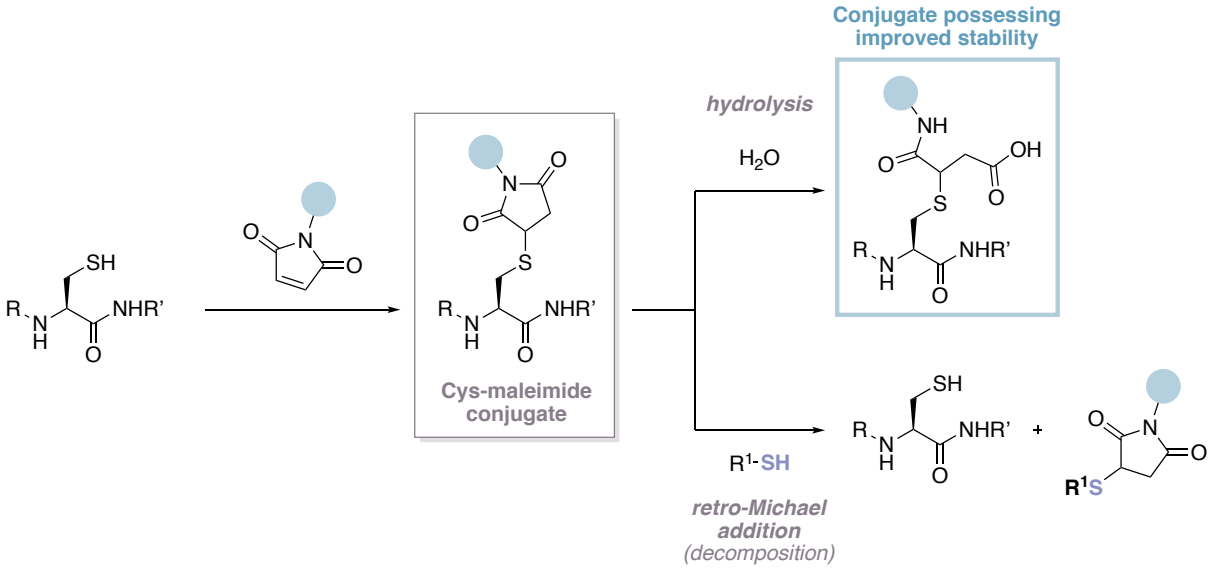
reagents (Scheme 5a),[22,39,40] aminoethylation (Scheme 5c),[41–44] Michael-type addition with vinyl sulfones[45–47] or carbonylacrylic reagents (Scheme 5d),[48] radical-based thiol-ene (Scheme 5g)[49–52] and thiol-yne chemistry (Scheme 5h)[52–56] and oxidative functionalisation to form mixed disulfides (Scheme 5f).[57–62]

Cysteine bioconjugation can also be achieved *via* disulfide rebridging (Scheme 5e). This involves reduction of a disulfide bond to form two free cysteine residues, followed by reaction with dielectrophilic reagents such as dibromomaleimide, which can reconnect the two cysteine residues.[63–67] One aim of the rebridging strategy is to avoid structural perturbation, which may be caused by breaking and functionalising the native disulfide bonds, through reforming unnaturally linked cysteine residues with exogenous functionality appended.



Scheme 5 Overview of traditional methods for cysteine bioconjugation

Despite its frequent use within chemical biology for cysteine bioconjugation, Michael addition to maleimides still faces some inherent limitations. The resulting conjugates tend to decompose in the presence of thiol nucleophiles or an external base *via* elimination of the maleimide through retro-Michael addition (Scheme 6).[68] This has prompted developments for the improved stability of maleimide-based conjugates *via* hydrolysis of the succinimide ring (Scheme 6),[69,70] but has also driven the expansion of the cysteine bioconjugation field with alternative reagents.

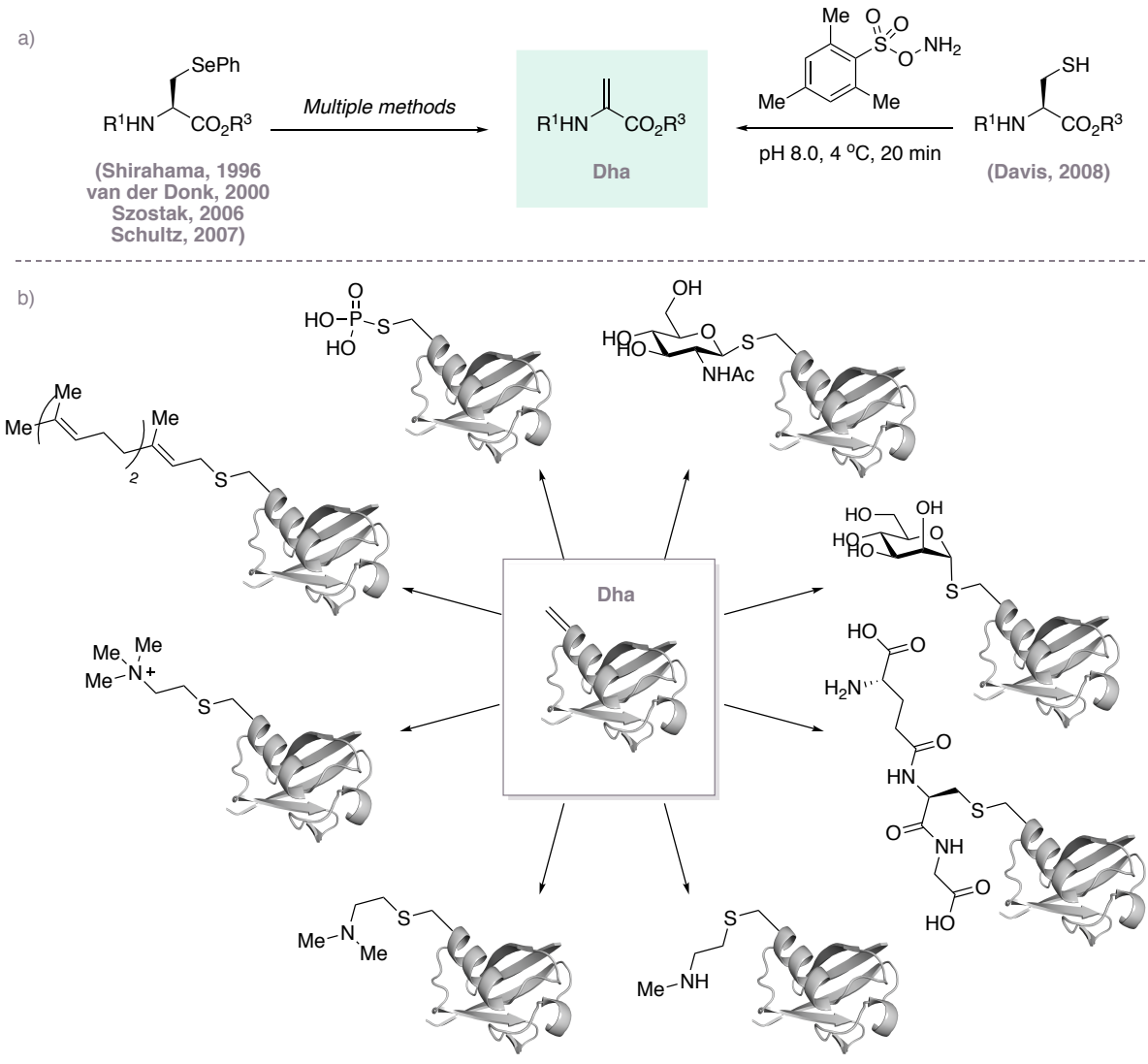


Scheme 6 Cysteine bioconjugation with maleimide reagents. Products undergo facile decomposition *via* retro-Michael addition of external thiol nucleophiles or an external base. Hydrolysis of the maleimide ring can lead to conjugates possessing improved stability.

1.2.2 Recent developments in cysteine functionalisation

A powerful protein functionalisation method pioneered by Davis and co-workers used cysteine residues to form the non-canonical amino acid dehydroalanine (Dha) which could provide a platform for further functionalisation (Scheme 7a).[71] Dha is a naturally occurring amino acid, although it is not encoded by DNA and instead arises through post-transcriptional modifications.[72] While the synthesis of Dha was reported as early as 1963 from serine,[73] its formation could also be carried out with the use of non-native alkylated selenocysteine[74] among other methods (Scheme 7a).[75–79] These earlier techniques could be applied to both peptides[77] and proteins,[78] but applications were limited to hydrolysis as a method of peptide digestion. While valuable, this did not constitute a bioconjugation method. However, advancements reported by Davis showed that Dha could be formed using milder conditions directly from

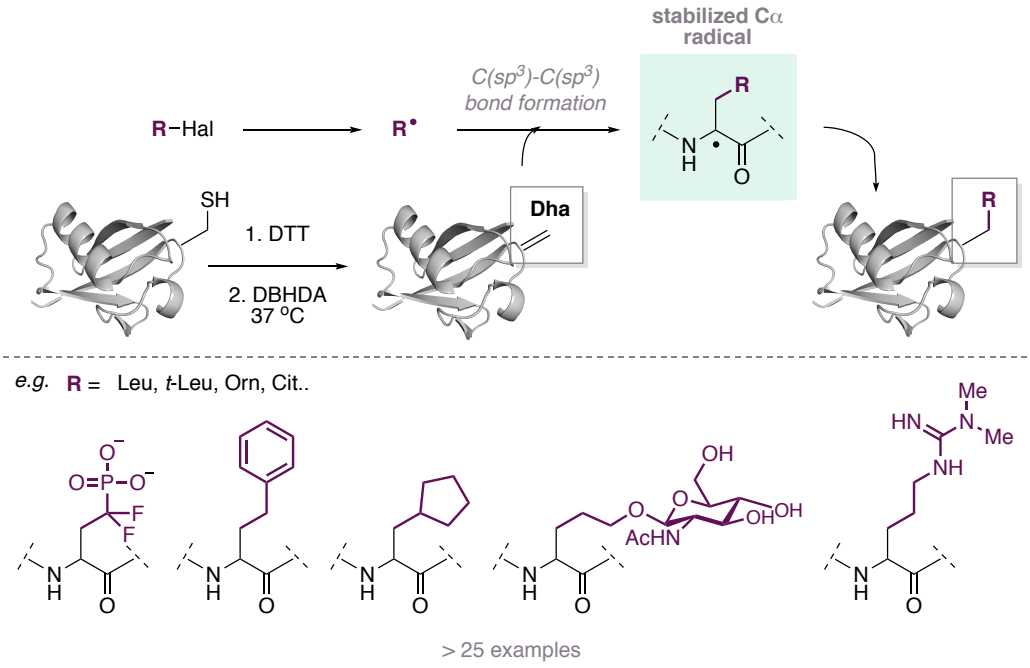
naturally occurring cysteine residues (Scheme 7a).[71] Subsequent Michael addition of Dha with varied thiols could install numerous functionalities, providing a novel bioconjugation approach used to mimic post-translational mutagenesis (Scheme 7b). However, the reaction proceeded with low diastereoselectivity,[80] and high concentrations of thiol-containing reagents were required which could potentially perturb native disulfide bonds forming mixed disulfides.[81] While technically a bioorthogonal transformation, this distinct 'umpolung' approach allows the typically nucleophilic cysteine residues to instead react with external nucleophiles in a two-step sequence.



Scheme 7 a) Strategies for the formation of dehydroalanine b) Dehydroalanine can react with a variety of thiols to form functionalised cysteine analogues which can mimic post-translational modifications[71]

Further work on Dha, reported by Davis, showed the chemical editing of amino acid residues *via* selective reaction at cysteine, mimicking posttranslational mutagenesis.[82] The authors reasoned that the use of mild, carbon-centred free-radical chemistry could prove more

biocompatible due to the tolerance of radicals to aqueous conditions and the unreactive nature of radicals to functionality present in biomolecules. Therefore, a remarkable radical-based strategy for the two-step bioorthogonal transformation of cysteine into alternative natural and non-natural amino acids was presented. Various alkyl halides could be used for the generation of primary, secondary and tertiary radicals (R•), which could react with Dha in peptides and proteins with excellent site- and regioselectivity (Scheme 8). However, the reaction displayed high sensitivity to oxygen, resulting in oxidative degradation through cleavage of the peptide backbone when not carried out under an inert atmosphere. Nonetheless, the method has since been applied to the α-deuteration of protein backbones as a potential tool for probing protein mechanisms.[83]

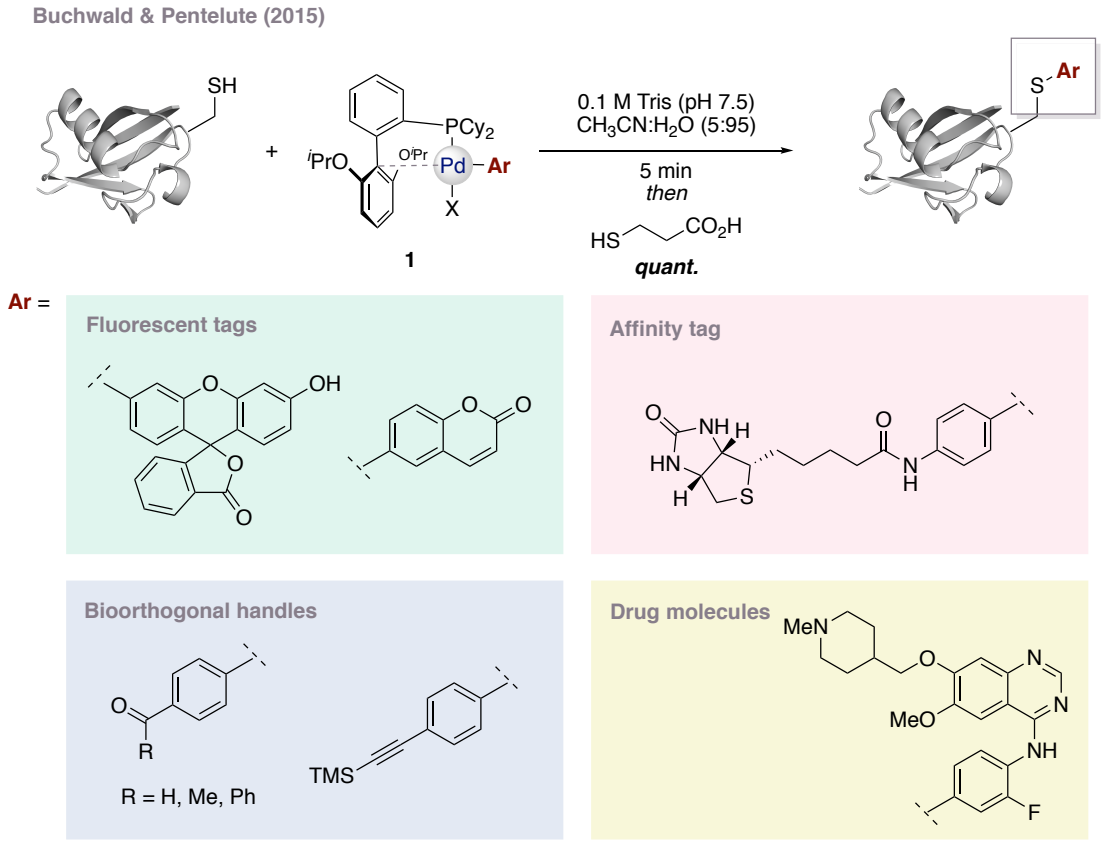


Scheme 8 Posttranslational mutagenesis through C(sp^3)–C(sp^3) bond forming radical reactions; DTT = dithiothreitol; DBHDA = 2,5-dibromohexanediamide; Orn = ornithine; Cit = citrulline

Owing to the fact that transition metal-catalysed reactions can be optimised to proceed with high chemoselectivity, mild conditions and excellent functional group tolerance,[84] there have been several reports of protein functionalisation using organometallic reagents.[84] These methods often require the use of functional handles[85] or non-natural amino acids (e.g. 4-iodo phenylalanine or aldehyde/alkyne containing amino acids),[84,86,87] providing a bioorthogonal approach. Additionally, high concentrations of reagents are generally required which could cause off-target reactivity and problems with purification.[88]

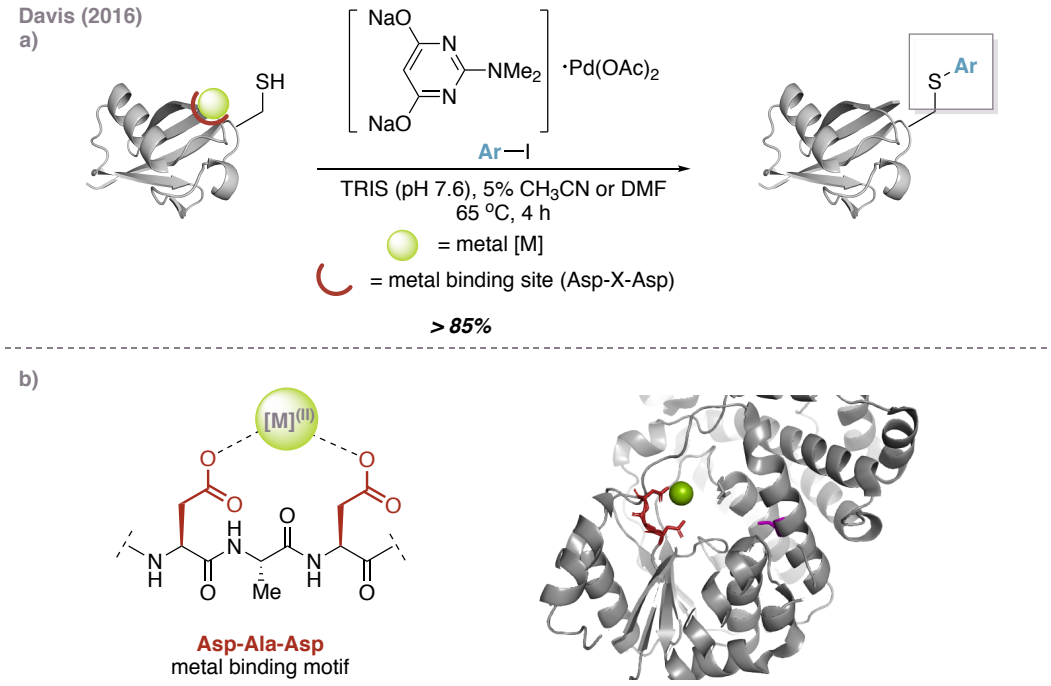
In light of these challenges, Buchwald, Pentelute and co-workers presented an elegant method for the bioconjugation of cysteine using organometallic palladium reagents (Scheme 9).[88] Using an aryl-palladium complex **1** with 2-dicyclohexylphosphino-2',6'-diisopropoxybiphenyl (RuPhos) as the ligand, quantitative conversion of cysteine was observed, yielding *S*-arylated products stable to acids, bases and external thiol nucleophiles. The arylation technique could be used to transfer a variety of biologically relevant aryl groups and was applied to the synthesis of antibody-drug conjugates. In contrast to other organometallic bioconjugation strategies targeting tryptophan and tyrosine, which require longer reaction times and achieve lower conversions,[84] the rate of the reaction was comparable to that of cysteine-maleimide conjugation even at low micromolar concentrations.

This cysteine arylation method was also applied to the synthesis of macrocyclic peptides, to ^{11}CN radiolabelling of unprotected peptides for *in vivo* imaging,[89] and to peptide and protein cross-linking.[90] Further developments additionally demonstrated the development of water soluble organometallic reagents, improving the biocompatibility of this approach.[91]



Scheme 9 Cysteine arylation using organometallic palladium reagents

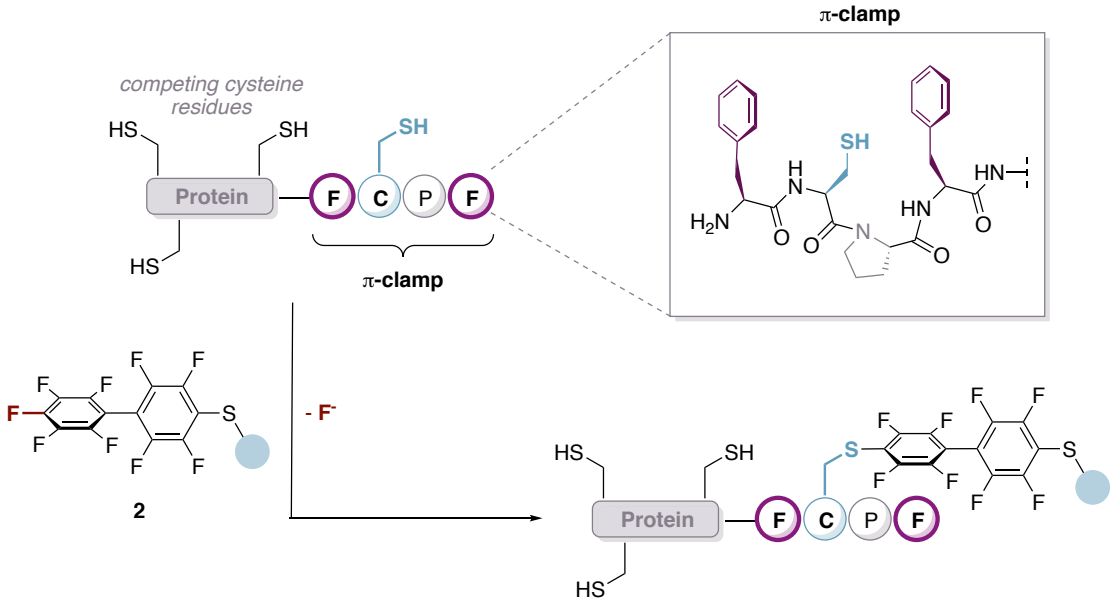
Davis and co-workers also reported a palladium-mediated cysteine arylation (Scheme 10).[92] A three amino acid recognition sequence (Asp-Ala-Asp) had previously been found to act as a natural metal binding motif and was exploited by Davis to direct a M(II)-aryl species, generated *in situ*, to the active site of a protein. Once in this active site, the M(II)-aryl species could react selectively with a proximal cysteine thiolate allowing high site-selectivity for cysteine arylation. Through mimicking the metal binding capabilities of natural protein systems and exploiting endogenous features to direct functionalisation, this proof-of-principle method represents a distinct strategy for selective bioconjugation. It could also be used for labelling, detection and identification of metal-dependent proteins in cell lysates.



Scheme 10 Cysteine arylation using organometallic palladium reagents exploiting an inhert binding motif within proteins a) General reaction scheme for cysteine arylation b) Schematic of metal binding to recognition unit within the protein mannosylglycerate synthase (visualisation created using Pymol, PDB = 2BO6) with Asp-Ala-Asp binding motif highlighted in red, naturally bound metal ion (here Mn^{2+}) highlighted in green and proximal Cys233 residue highlighted in pink

Through a similar exploitation of a recognition motif, Pentelute and coworkers reported the selective perfluoroarylation of single cysteine residues within proteins using nucleophilic aromatic substitution (Scheme 11).[93] The serendipitously discovered "π-clamp", a four amino acid sequence (Phe-Cys-Pro-Phe), was reported to adjust the microenvironment of the clamp's cysteine residue within a complex peptide or protein. Structural and computational mechanistic studies confirmed that the role of the proline was to position the residues into a conformation which promoted the reaction, highlighting the significance of the *trans* conformation prolyl amide bond of the π-clamp.[94] Additionally, distinct interactions between the perfluoroaryl

electrophile **2** and the phenylalanine side chain were observed, with the phenylalanine residues also thought to be activating the cysteine thiol.[94] In contrast to the work from Davis (Scheme 10) which could exploit a recognition unit inherent to the biological system, incorporation of the π-clamp motif into the biomolecule of interest was required prior to cysteine bioconjugation.

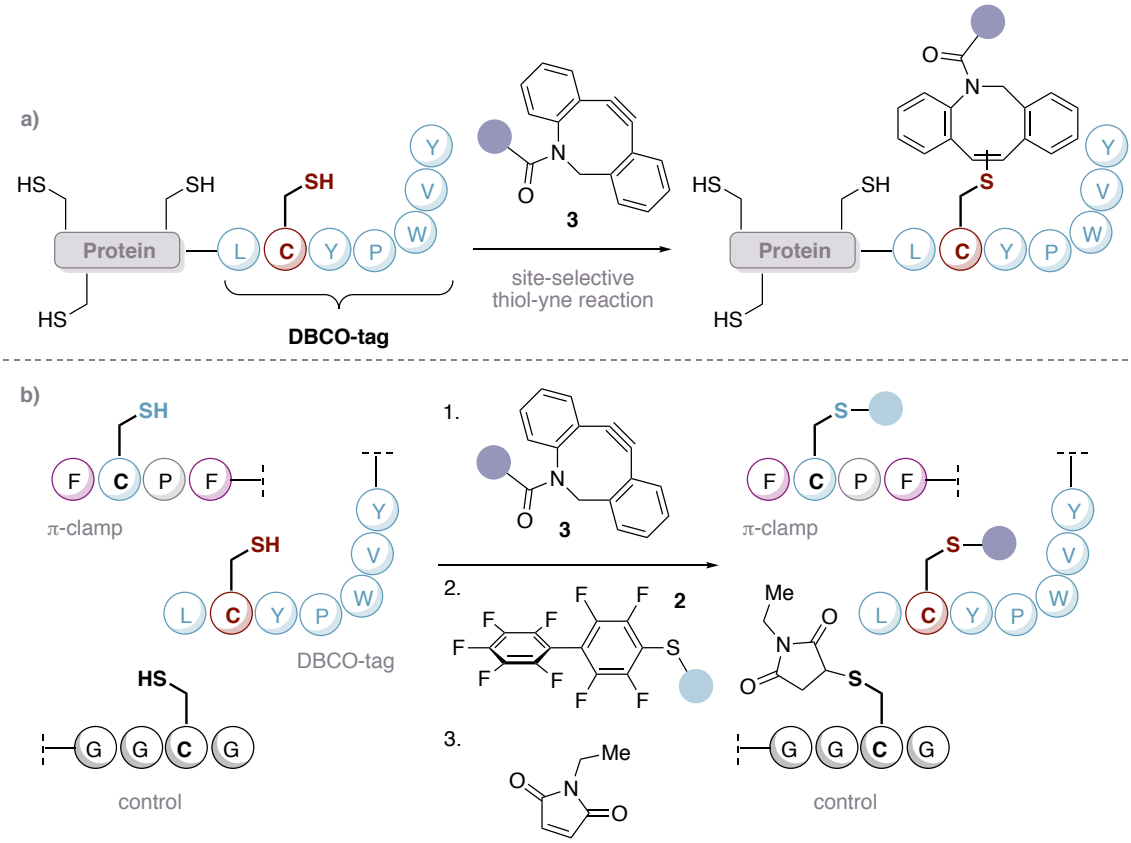


Scheme 11 Selective cysteine conjugation using a 4 amino acid π-clamp sequence (Phe-Cys-Pro-Phe) for selective perfluoroarylation

The authors have since shown that the concept could be applied to a longer, 29 amino acid abiotic peptide.[95] Although this strategy requires the pre-installation of the π-clamp, its utility was demonstrated through the covalent attachment of the π-clamp sequence to an enzyme, allowing regiospecific, single-site perfluoroarylation. Most importantly, regioselective perfluoroarylation was achieved whilst leaving the enzyme's active-site cysteine residue unchanged.[95]

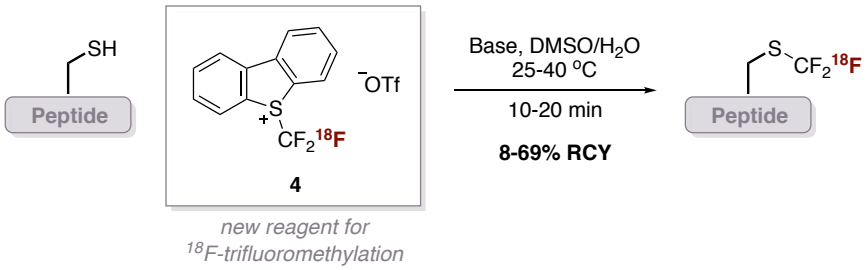
Recently, Pentelute *et al.* have also applied a similar concept to thiol-yne chemistry for cysteine-cyclooctyne bioconjugation, developing a peptide recognition unit which could react in a sequence-selective fashion with aza-dibenzocyclooctyne (DBCO) reagents **3** (Scheme 12a).[96] Named a "DBCO-tag", this seven amino acid sequence facilitates the use of commercially available DBCO reagents for regioselective modification of single cysteine residues on proteins which contain multiple endogenous cysteines. While the exact structural and mechanistic role of this recently developed DBCO-tag is currently not well understood, it has enabled the transfer of multiple functional payloads to proteins and has been used for the site-selective labelling of antibodies. Application of this method to a mixture of polypeptides allowed an

unprecedented site-selective labelling of three unprotected cysteine-containing peptides in a single solution using three different reagents (Scheme 12b).[96] This exploited the π-clamp sequence, the DBCO-tag and Michael addition to maleimide reagents to sequentially and selectively label three cysteine residues with unique microenvironments within a complex mixture.



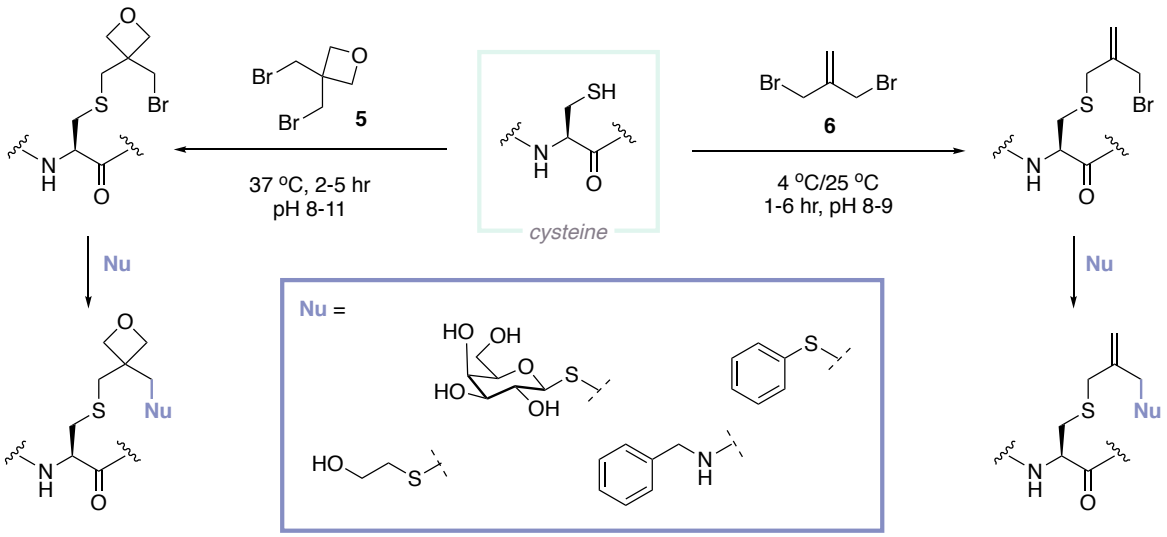
Scheme 12 a) Site-selective cysteine-cyclooctyne conjugation b) One-pot sequential labelling of three unprotected cysteine-containing peptides combining DBCO-tag, π-clamp and maleimide conjugation methodologies

The radiolabelling of peptides has received considerable attention owing to its application in positron emission tomography (PET) for the visualisation of biochemical processes *in vivo*. PET techniques involve the introduction of a positron-emitting radioisotope such as ^{11}C or ^{18}F. However, strategies for the radiolabelling of peptides commonly rely on bioorthogonal techniques, requiring pre-functionalisation of the peptide.[97] In 2018, Davis and Gouverneur described the first protocol enabling direct ^{18}F-labelling of cysteine residues in unmodified peptides. They use a newly developed ^{18}F-Umemoto reagent **4** to introduce a trifluoromethyl group (Scheme 13).[98] Biodistribution studies revealed the [^{18}F]-SCF_3 moiety was viable for *in vivo* imaging, however mixed DMSO/water solvent systems were used in the reaction, potentially limiting its applicability towards larger protein systems.



Scheme 13 ^{18}F labelling of unmodified peptides with a CF_3 group; RCY, radiochemical yield

Bernardes and co-workers reported approaches to enable the introduction of an electrophilic handle at cysteine residues (Scheme 14).[99,100] This approach was similar to the work of Davis for the conversion of cysteine to Dha and its subsequent reaction with external nucleophiles, providing an analogous 'umpolung' strategy. The site-selective reaction with homobifunctional electrophilic reagents **5** or **6** could install an electrophilic handle at cysteine residues. Now, the inherently nucleophilic cysteine residue could instead undergo subsequent reaction with various external nucleophilic groups, providing a complementary strategy for bioorthogonal transformations at cysteine. However, the requirement for basic pH presented a limitation of the method when translating to biological systems.



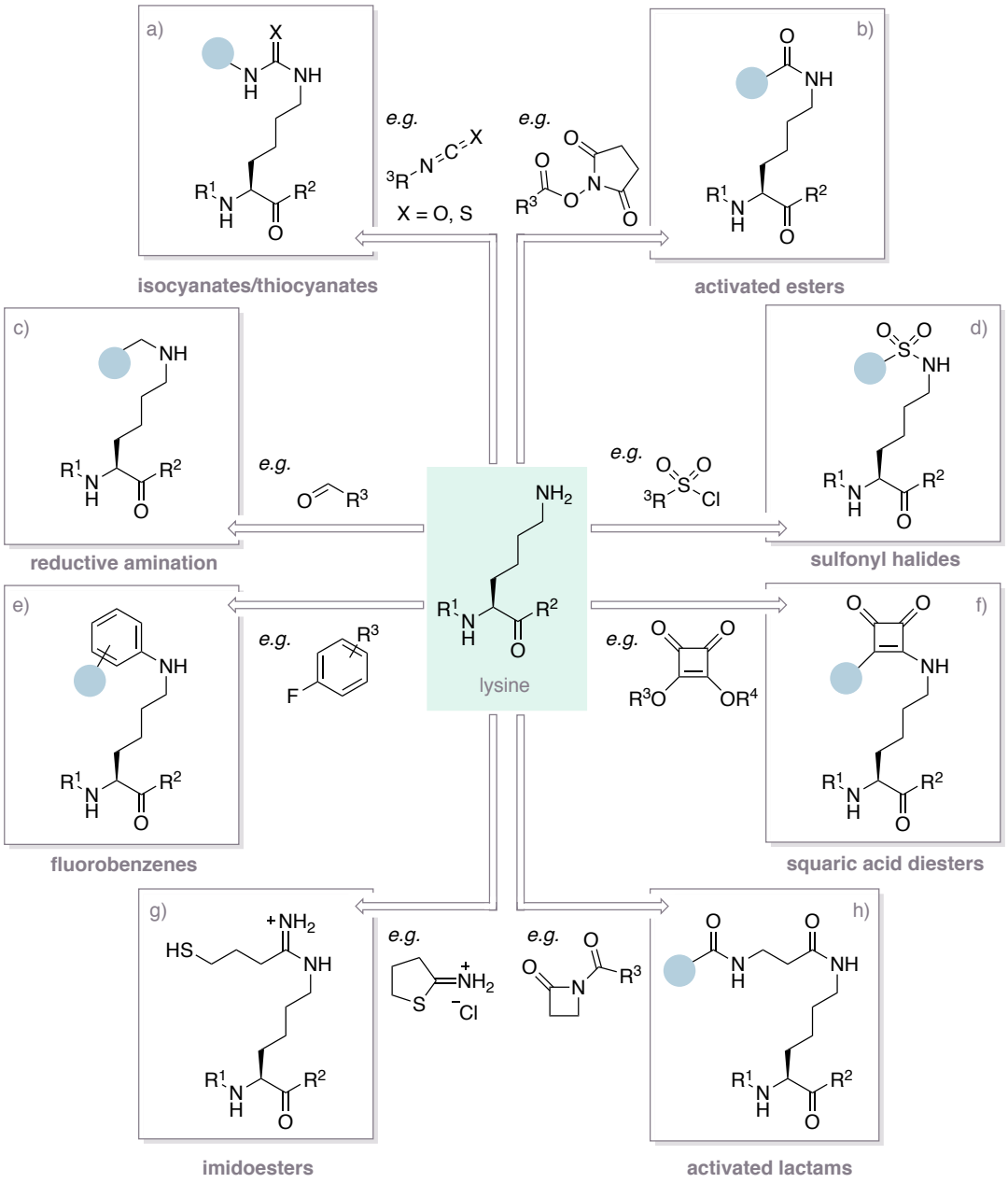
Scheme 14 Site-selective installation of an electrophilic handle at cysteine residues for subsequent reaction with external nucleophiles

1.2.3 Bioconjugation of lysine

In addition to cysteine, lysine is also an important target for bioconjugation. Lysine residues are often exploited due to their well-established reactivity and the abundance of methods for the modification of primary amines. In contrast to cysteine (natural abundance ~3% in vertebrates), lysine is one of the most common amino acids (~7% in vertebrates),[23] with 20 or more lysine residues frequently present on the surface of a protein.[3] Owing to its high natural abundance, bioconjugation at lysine residues often results in heterogeneous mixtures of products with varying degrees of conjugation, meaning strategies are seldom regioselective. However, when multiple conjugations are desired, or when selectivity is less important, lysine bioconjugation has found broad utility.[3]

With a pKa of approximately 10,[2] lysine residues are protonated under biological conditions. In this state, lysine possesses no significant nucleophilicity and most amine-reactive reagents will instead modify any accessible cysteine residues due to the superior nucleophilicity of the sulfhydryl side chain. To circumvent this, it is important to maintain basic conditions (pH > 8) to obtain optimal yields and selectivity for bioconjugation.[81]

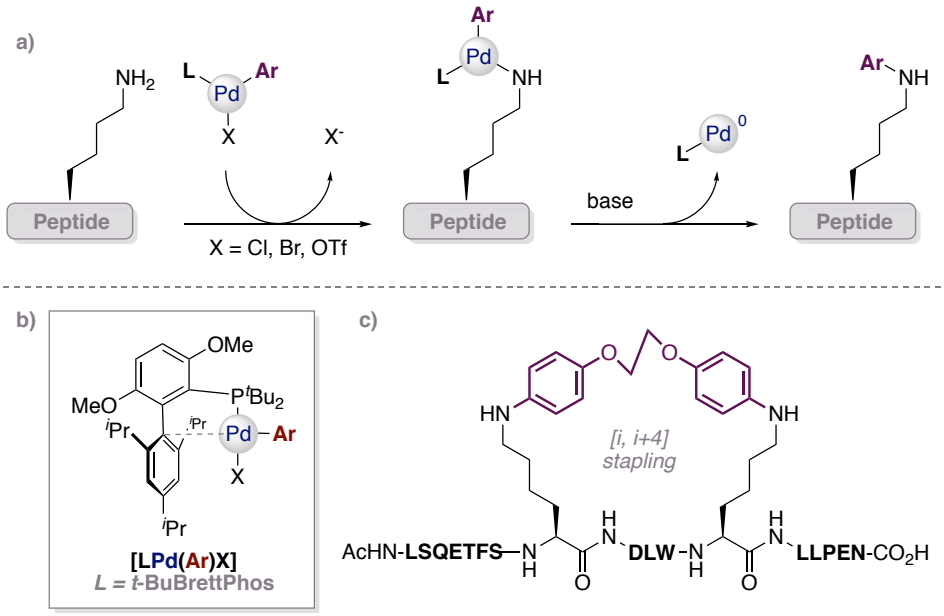
Scheme 15 provides an overview of strategies for the functionalisation of lysine. Strategies involving activated esters (*N*-hydroxysuccinimide (NHS)[101] and sulfo-NHS,[102] Scheme 15b) and imido esters[103–105] (Scheme 15g) have proven the most applicable, and many of these compounds have been developed into commercially available reagents for bioconjugation.[20] However, depletion of these highly reactive reagents can occur in water *via* hydrolysis.[2] Alternative methods have employed isocyanates[106] or isothiocyanates (Scheme 15a),[107–109] sulfonyl halides (Scheme 15d),[110–113] squaric acid diesters (Scheme 15f),[114,115] activated lactams (Scheme 15h),[116–118] fluorobenzenes (Scheme 15e)[119–121] or reductive amination with aldehydes (Scheme 15c).[122–125] These will not be discussed in detail and the reader is directed to a number of extensive reviews.[2,126–128]



Scheme 15 Overview of established methods for the bioconjugation of lysine involving the reaction of lysine with various electrophilic reagents

Following on from their work on palladium-mediated cysteine arylation,[88] a recent example from Buchwald and Pentelute aimed to expand their strategy for the arylation of lysine. They reported a general lysine arylation method, to form stable *N*-aryl conjugates using preformed or *in situ* generated [LPd(Ar)X] complexes (Scheme 16a,b).[129] The biaryl phosphine ligand, in combination with mildly basic sodium phenoxide, enabled the functionalisation of complex peptide substrates and found application in the synthesis of macrocyclic peptides (Scheme 16c). While the protocol was shown to be completely selective in the presence of most other nucleophilic residues, cysteine was not tolerated due to competitive base-mediated

dehydroalanine formation. Undesirable diarylation was observed in the case of the guanidinium side chain of the amino acid arginine.

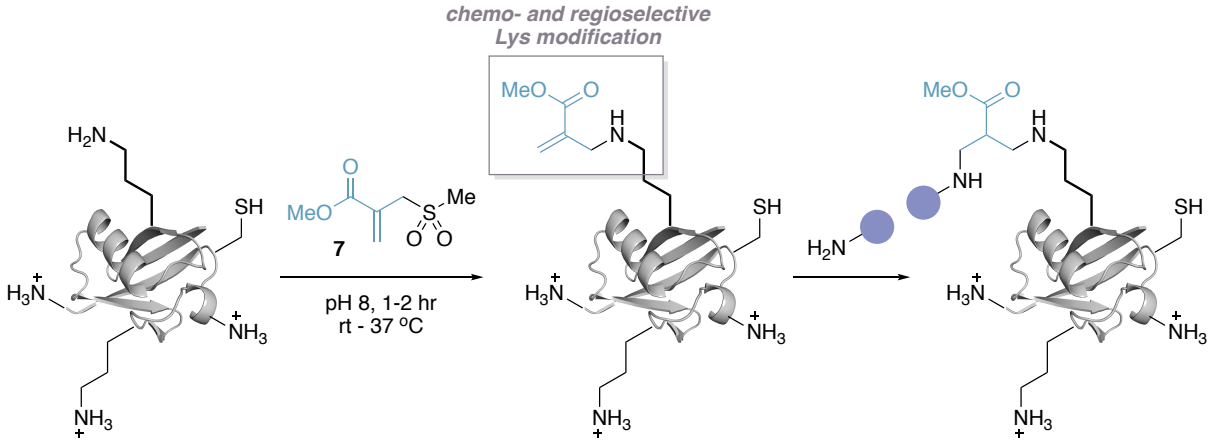


Scheme 16 a) Palladium-mediated arylation of lysine in unprotected peptides b) Most effective [LPd(Ar)X] complex, using ligand *t*-BuBrettPhos c) Macrocyclic peptide formed using polymetalated Pd complex

In light of the challenges associated with more traditional lysine bioconjugation strategies, namely low regioselectivity and instability of reagents,[2] recent approaches have sought to circumvent these issues. Despite the abundance of methods for the functionalisation of lysine, a strategy which could label a single lysine residue in a protein containing multiple lysines has remained elusive. It is understood that variations in the local microenvironment of amino acids can induce innate and subtle reactivity differences between different amino acids bearing the same side chain within a protein.[130] In the case of lysine, the pKa of the amine side chain can vary by up to 5 units depending on the polarity of its microenvironment, its solvent accessibility, protein folding and the nature of the surrounding residues.[131] This difference in pKa can incite distinct reactivity, with some residues proving 'hyper-reactive' towards certain electrophiles. Strategies attempting to exploit hyper-reactive lysine residues have been reported, although they remain poorly understood and largely substrate dependent.[132–134]

In 2017, Bernardes *et al.* pioneered the use of computer-aided design to exploit the varying pKa of lysine residues within proteins for both chemo- and regioselective modification of a single, specific lysine (Scheme 17).[130] At slightly basic pH, the lysine residue with the lowest pKa was kinetically favoured for reaction with sulfonyl acrylate reagents **7**. Upon lysine labelling, spontaneous elimination of methanesulfinic acid generated an acrylate electrophile which could

undergo subsequent aza-Michael addition with an external amine reagent. The authors proposed computational evidence that the sulfonyl acrylate reagents were selective for lysine over other nucleophilic groups as a result of a transient hydrogen bond between the sulfonyl group and the ε-amino group of lysine which stabilised the intermediate formed. This method highlights the power of precise prediction of a residue's microenvironment within a protein, facilitating rational design of a selective bioconjugation approach.



Scheme 17 Chemo- and regio-selective lysine modification on a native protein using computer-designed sulfonyl acrylate reagents

Cysteine and lysine residues remain the most widely explored amino acids and provide platforms for state-of-the-art bioconjugation techniques. However, approaches targeting these amino acids still present some inherent limitations. Many strategies remain sequence specific and, in the case of cysteine, require pre-installation of recognition units or incorporation of free cysteine residues to achieve high selectivity. In order to address these limitations, attention has focused on the targeting of alternative native functionality within proteins. Through the development of distinct bioconjugation strategies, it may be possible to not only overcome issues fundamental to cysteine and lysine bioconjugation, but may also provide a complementary strategy to enable the formation of multi-functionalised protein scaffolds.[19] Formation of multi-functionalised conjugates through mutually orthogonal reactions used in tandem are desirable due to the potential application of the resulting products in simultaneous metabolic imaging of two different processes or in monitoring protein dynamics.[19]

To this end, strategies for the selective functionalisation of several other native functional groups have been reported, including the bioconjugation of tryptophan,[135–139] histidine,[140–142] tyrosine[143–147] and the N-terminus[148–151] and C-terminus[152] (Figure 2). These methods have been extensively reviewed elsewhere.[2,4] Ongoing work aims to expand the repertoire of methods

available for targeting these alternative native motifs for bioconjugation reactions of polypeptides and proteins.

One particular amino acid of interest for the development of a novel bioconjugation strategy is methionine (Figure 2). Possessing relatively low natural abundance (~2% in vertebrates)[23] and simple ancillary protein function,[81] targeting of this amino acid presents a potential opportunity for addressing the limitations inherent to cysteine and lysine bioconjugation.

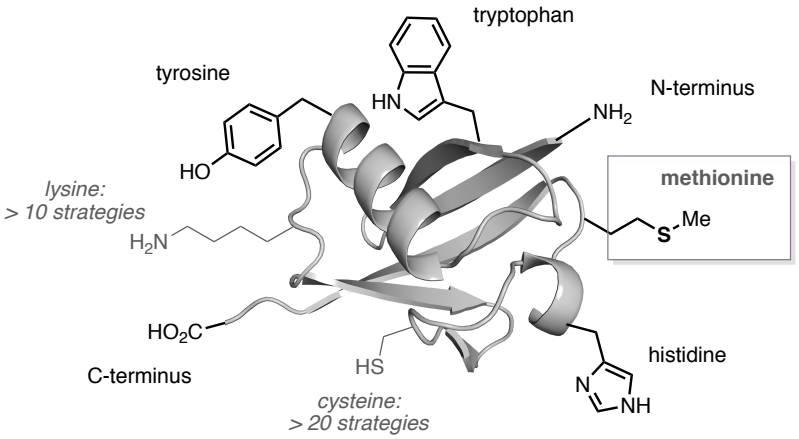


Figure 2 Figure highlighting the alternative amino acid side chains and the C- and N-terminus which have been targeted for distinct bioconjugation strategies. Methionine is highlighted as a potentially favourable target for bioconjugation

1.3 Methionine bioconjugation

Methionine is encoded by the AUG start codon at the beginning of protein synthesis, but is often post-translationally excised, resulting in low natural abundance (~2%).[23] Bearing a hydrophobic thioether side chain which is commonly buried within a protein's hydrophobic core, solvent exposed methionine residues are rarely found on the surface of proteins.[81] Therefore, it may be possible to target single methionine residues to provide an effective strategy for the formation of homogeneous, well-defined protein conjugates. The polar, non-ionisable side chain is the only amino acid residue which remains nucleophilic at low pH, thereby offering complementary reactivity to other polar, ionisable amino acids.[126]

Additionally, within most proteins, methionine plays a simple structural role, with a third of all proteins found to be stabilised by S/π interactions between the sulfur atom and other aromatic amino acids. Methionine residues involved in these stabilising interactions are generally found in the hydrophobic core of proteins and hence are not solvent exposed and are unavailable for

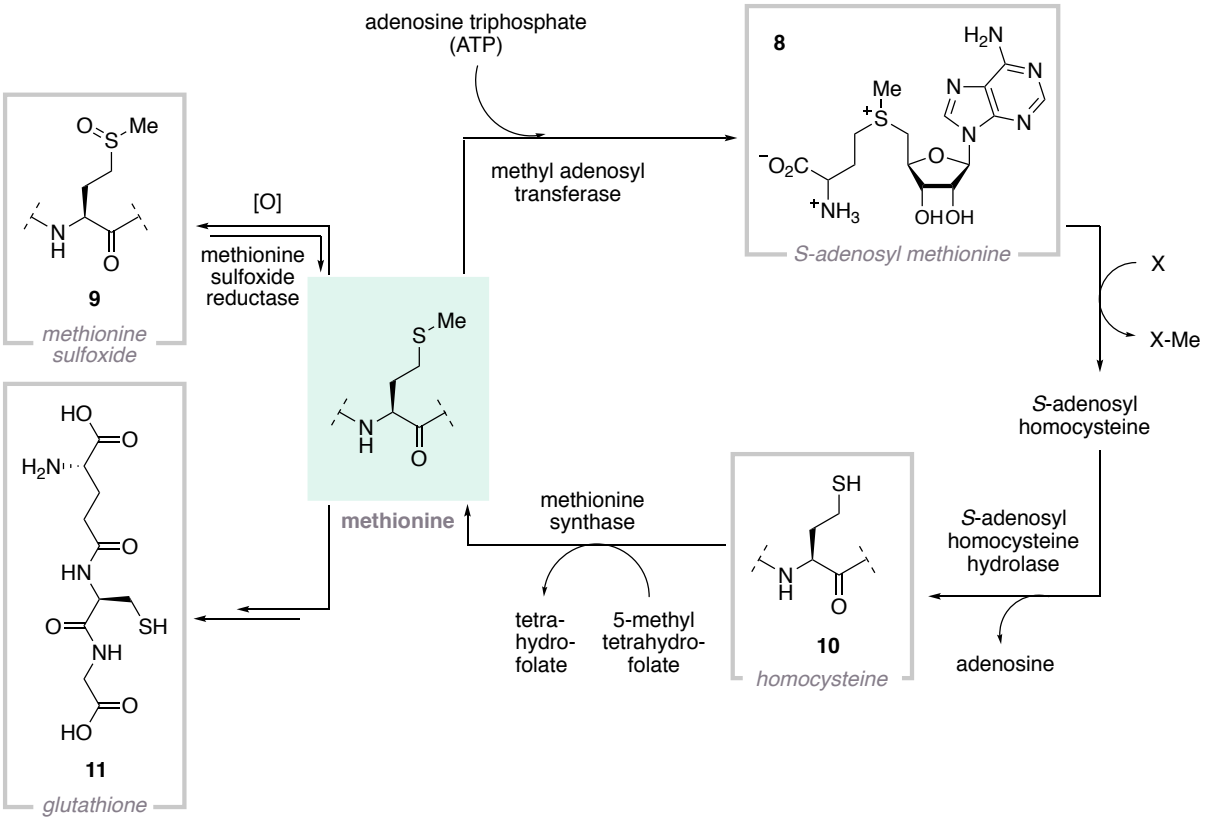
bioconjugation.[153] The replacement of methionine with norleucine, an amino acid with a straight hydrocarbon chain, showed no detrimental effect on protein structure or function, demonstrating methionine's ancillary role.[154]

Methionine is the key component of the cofactor *S*-adenosyl methionine **8** (SAM). As the main cellular carrier of methyl groups within the body, SAM plays an important role alongside a SAM-dependent methyl transferase enzyme to transfer methyl groups to various biomolecules including DNA.[155,156] However, SAM is naturally biosynthesised from *L*-methionine in living cells and is not a target for bioconjugation reactions.

The relatively high oxidation potential of the methyl thioether side chain of methionine[157] lends itself to reversible, biochemical oxidation processes which can be exploited by biological systems to protect against oxidative stress by reactive oxygen species.[158] Proteins that are damaged by oxidative stress are often targeted for proteolysis, or other protein degradation pathways, to prevent cellular damage.[159] To avoid this degradation an antioxidant enzyme, methionine sulfoxide reductase, can catalyse the reduction of methionine-*S*-sulfoxide **9** back to methionine (Scheme 18), which is no longer targeted for proteolysis.[159] Therefore, methionine plays a regulatory role in endogenous antioxidant enzymes and key metabolic processes and in some proteins methionine can also function as a redox sensor.[160]

As oxidative stress has been linked to pathological diseases such as cancer, harnessing the redox properties of methionine may assist in reducing the prevalence of cancer.[161,162] The inhibition of cancer cell growth has been demonstrated through the metabolism of methionine to homocysteine **10**,[163] while methionine has also been shown to chelate lead and remove it from tissues, decreasing oxidative stress.[164] Methionine can also directly influence the functioning of the immune system through the production of the low molecular weight antioxidant glutathione **11**.[158]

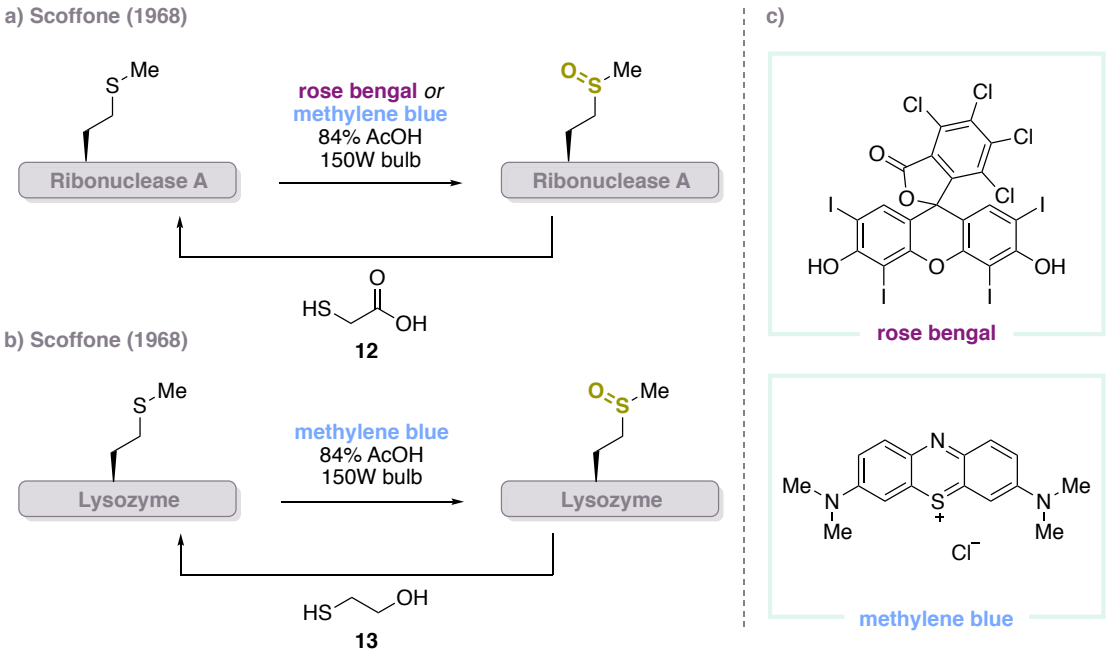
The oxidation and reduction of methionine is a reversible process, repeatedly modifying and regenerating the thioether side chain in response to external stimuli. This redox process means that the thioether of methionine may still be available for bioconjugation even after being exposed to an oxidising environment within biological systems.



Scheme 18 Methionine and the synthesis/metabolism of its derivatives within biological systems

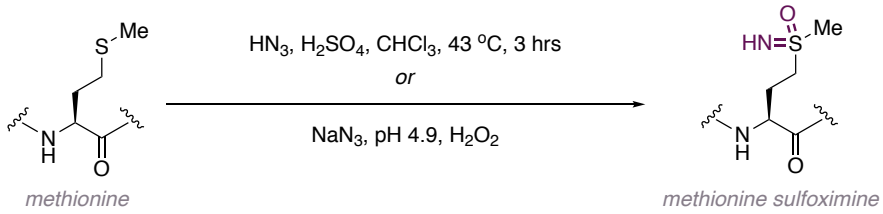
The low natural abundance and ancillary function of methionine presents the opportunity to functionalise proteins without perturbing protein structure and function. Additionally, selective labelling could also allow the investigation into the unique biological roles of methionine within more specific systems. Therefore, a strategy to selectively target this amino acid could prove invaluable. Previously reported strategies for the selective functionalisation of methionine are comparable to its biochemical roles, with the methyl thioether providing a platform for oxidation to the sulfoxide, or for alkylation to sulfonium salts.

Inspired by the redox properties of the thioether of methionine, Scoffone demonstrated the selective, dye-sensitised photooxidation of methionine in ribonuclease A[165] (Scheme 19a) and lysozyme[166] (Scheme 19b). Using the photosensitisers rose bengal or methylene blue (Scheme 19c), quantitative conversion to methionine sulfoxide was observed, causing an 87% decrease in enzymatic activity in the case of ribonuclease A, due to the resulting change in conformation. In both cases, oxidation was found to be reversible, using either thioglycolic acid **12** or 2-mercaptoethanol **13** as a reductant, restoring full enzymatic activity on regeneration of the native methionine thioether.



Scheme 19 Reversible, dye-sensitised photooxidation of a) Ribonuclease A and b) Lysozyme. c) photo-sensitisers

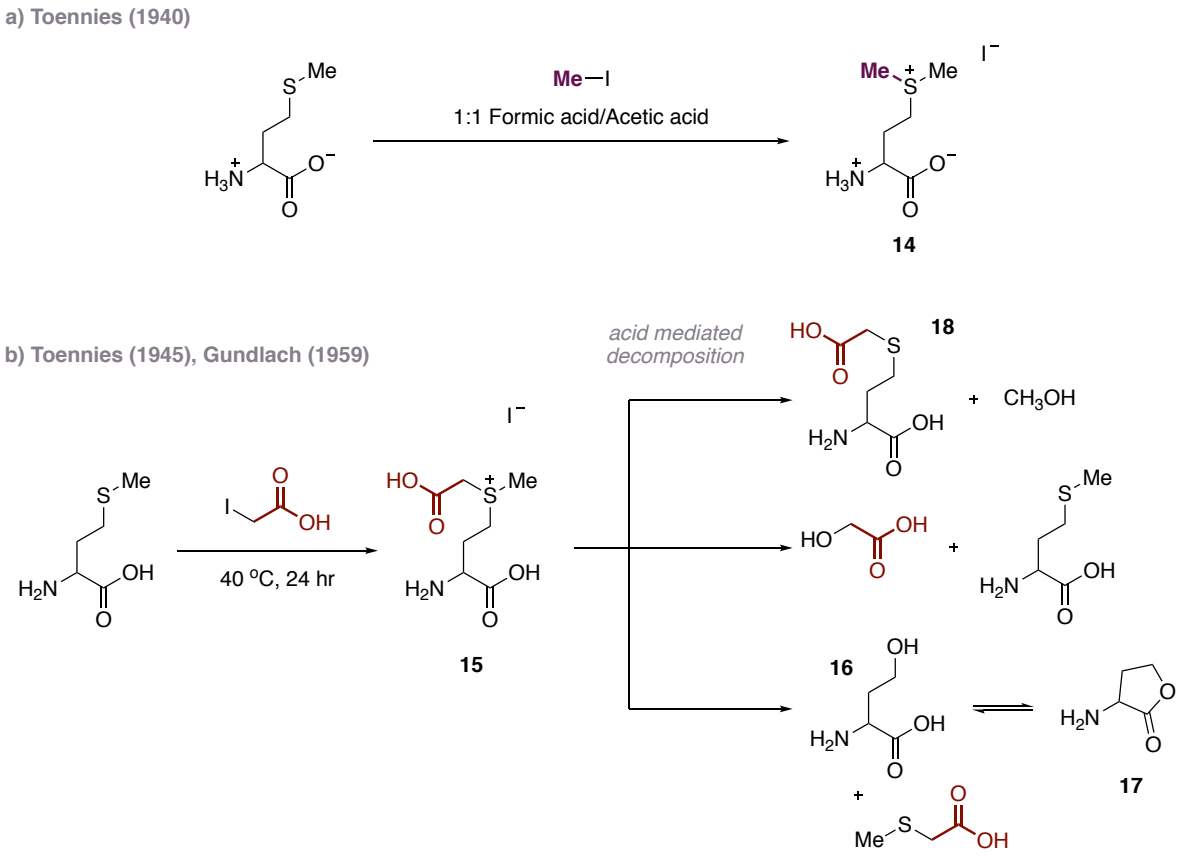
In a similar oxidative strategy, Whitehead and Bentley reported that the thioether of methionine could be converted to a sulfoximine using sodium azide or hydrazoic acid (Scheme 20).[167,168] This technique was applied to the functionalisation and inhibition of the enzyme horseradish peroxidase and provided evidence for the presence of a methionine residue in its active site.[169] However, while both of these oxidative strategies probe enzymatic function, they employ strongly acidic conditions which are largely incompatible with most biological systems.



Scheme 20 Oxidative functionalisation of methionine with azide reagents

Arguably, compared to these oxidative approaches, more impactful discoveries in the field of protein functionalisation have resulted from the development of alkylative strategies targeting methionine, as these strategies have presented the opportunity to incorporate a functional handle or payload. As early as 1940, Toennies reported pioneering work for the functionalisation of methionine using methyl iodide to form the corresponding *S*-methyl sulfonium species **14** (Scheme 21a).[170] This could be applied to the enzyme ribonuclease A as a method for the identification of the presence of surface exposed methionine residues.[171] Alternative alkylation

strategies emerged from 1945,[172] reported on simple systems by Toennies and subsequently by Gundlach in 1959.[173] Both described the reaction of methionine with α-halo acids forming carboxymethyl sulfonium salts **15** (Scheme 21b). While α-halo acids have been previously shown to react with free cysteine residues, in the absence of cysteine and under acidic conditions these reagents can selectively react with methionine. Exposing the resulting sulfonium salt to acidic conditions led to decomposition, yielding homoserine **16** and the corresponding homoserine lactone **17** through loss of a neutral sulfide (HO_2CCH_2SMe). Alternative degradation pathways yielded *S*-carboxymethyl homocysteine **18** and regenerated methionine through dealkylation.[173]



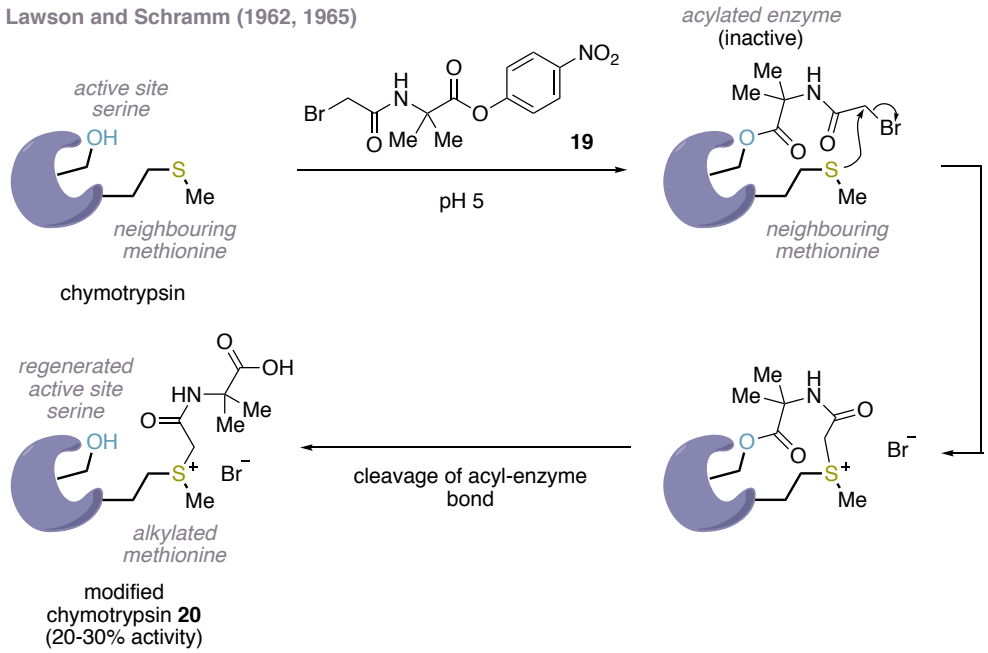
Scheme 21 Early work for the alkylation of methionine using methyl iodide and α-halo acids

Aside from the ancillary biological role and redox properties, there are specific examples in which methionine plays a more important role within proteins, namely it can be found as the key residue in the active site of many kinases.[174] In these situations the functionalisation of methionine may be a valuable tool to allow investigation into its more complex biological roles.

To this end, Gundlach applied the use of α-halo acids to the inhibition of the enzyme ribonuclease, known for its robust nature.[175] Labelling at methionine could only be confirmed

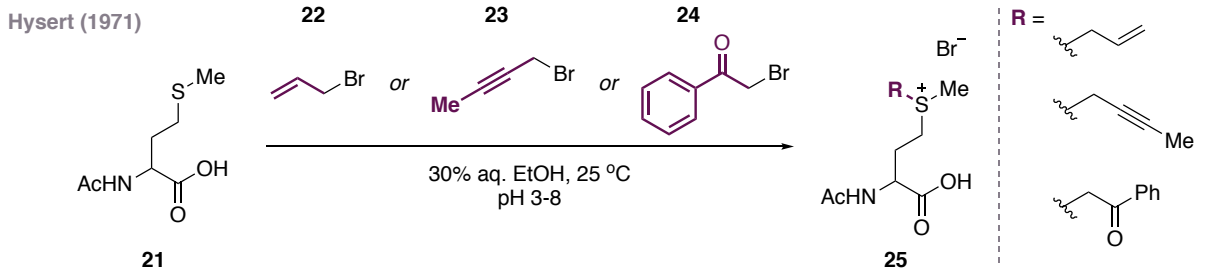
through detection of degradation products resulting from functionalisation of the thioether side chain. Additionally, depending on the reaction conditions, the modification of other amino acids such as lysine and histidine was observed. The use of α-halo acids was also applied to the functionalisation of the enzyme fumarase in 1969 using iodoacetate,[176] however again the process lacked selectivity, yielding a mixture of products resulting from reaction with histidine residues in addition to methionine.

Lawson and Schramm demonstrated that more complex α-halo carbonyl derivatives could be used to label a methionine residue adjacent to the active site of the enzyme chymotrypsin (Scheme 22).[177,178] Through the use of *p*-nitrophenyl bromoacetyl-α-aminobutyrate **19** at pH 5, a serine residue in the active site of the enzyme was acylated. A subsequent, proximity-driven alkylation of a neighbouring methionine residue could occur, followed by cleavage of the acyl-enzyme bond. While modification of the serine residue caused full deactivation of the enzyme, alkylation of methionine partially restored the enzymatic activity (20-30% activity of standard chymotrypsin). Restoration of the enzymatic activity was largely due to regeneration of the free serine residue, however the reduced activity of the trialkyl sulfonium enzyme **20** indicated that methionine may play a role within the active site. The authors propose that this bioconjugation strategy could prove a useful tool for the mapping of enzyme active sites.



Scheme 22 Inhibition of chymotrypsin through alkylation of methionine; labelling of an active site serine residue using a bifunctional reagent allowed proximity-driven alkylation of a methionine residue followed by cleavage of the acyl-enzyme bond

Developments from Hysert *et al.* expanded the scope of reagents for the direct alkylation of methionine.[179] They employed allyl bromide **22**, 1-bromo-2-butyne **23** or 2-bromoacetophenone **24** with *N*-acetyl-DL-methionine **21** at pH 3–8 to form the corresponding sulfonium species **25** (Scheme 23). Importantly, no decomposition of the sulfonium products was observed, indicating the potential stability of these sulfonium motifs. A drawback of this method was the requirement for organic solvent in order to facilitate methionine modification. Additionally, the reaction was only able to achieve full conversion after 24 hours.



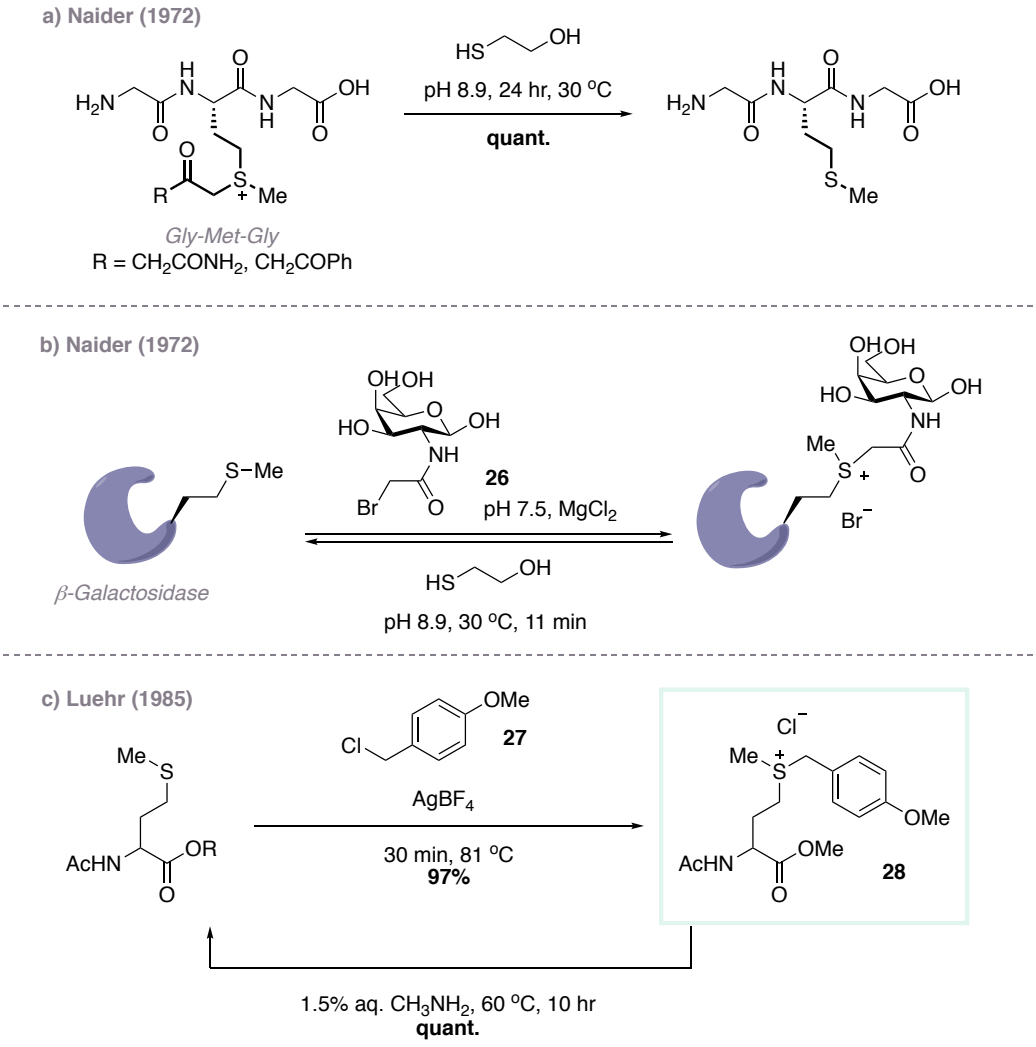
Scheme 23 Broader scope of alkylating reagents for alkylation of methionine

Following the alkylation of methionine residues with α-halo carbonyl reagents, it was reported that dealkylation of the resulting sulfonium species could also be achieved, in many cases restoring enzymatic activity. Naider showed that sulfur nucleophiles could regenerate methionine residues from their sulfonium salts on simple peptidic systems (Scheme 24a).[180] Although reaction times were long and required elevated temperatures, this reversible alkylation of methionine was applied to a methionine residue close to the active site of the enzyme β-galactosidase (Scheme 24b).[181] It was hypothesised that the active site could bind the alkylating agent, bringing this into close proximity to methionine for selective functionalisation. Using kinetic studies, it was concluded that the inactivation of β-galactosidase using *N*-bromoacetyl-β-D-galactosylamine **26** did proceed *via* the formation of a reversible enzyme-substrate complex prior to covalent bond formation, with bond formation suggested to be directed by the enzyme's active site. This proved distinct to alkylation with iodoacetamide which obeyed a second-order rate law, indicating a simple bimolecular reaction.

A reversible alkylation procedure was also described by Luehr in which *p*-methoxybenzyl chloride **27** and silver tetrafluoroborate in acetonitrile could afford the *S*-benzyl sulfonium species **28** in just 30 minutes from protected methionine amino acid residues (Scheme 24c).[182] The use of silver tetrafluoroborate allowed alkylation using unactivated alkyl halides. Heating this sulfonium compound in aqueous methylamine regenerated the methionine thioether

quantitatively, however complete hydrolysis of the methyl ester protecting group was observed. In this example, the forcing conditions required for dealkylation would likely prove incompatible with larger biomolecules.

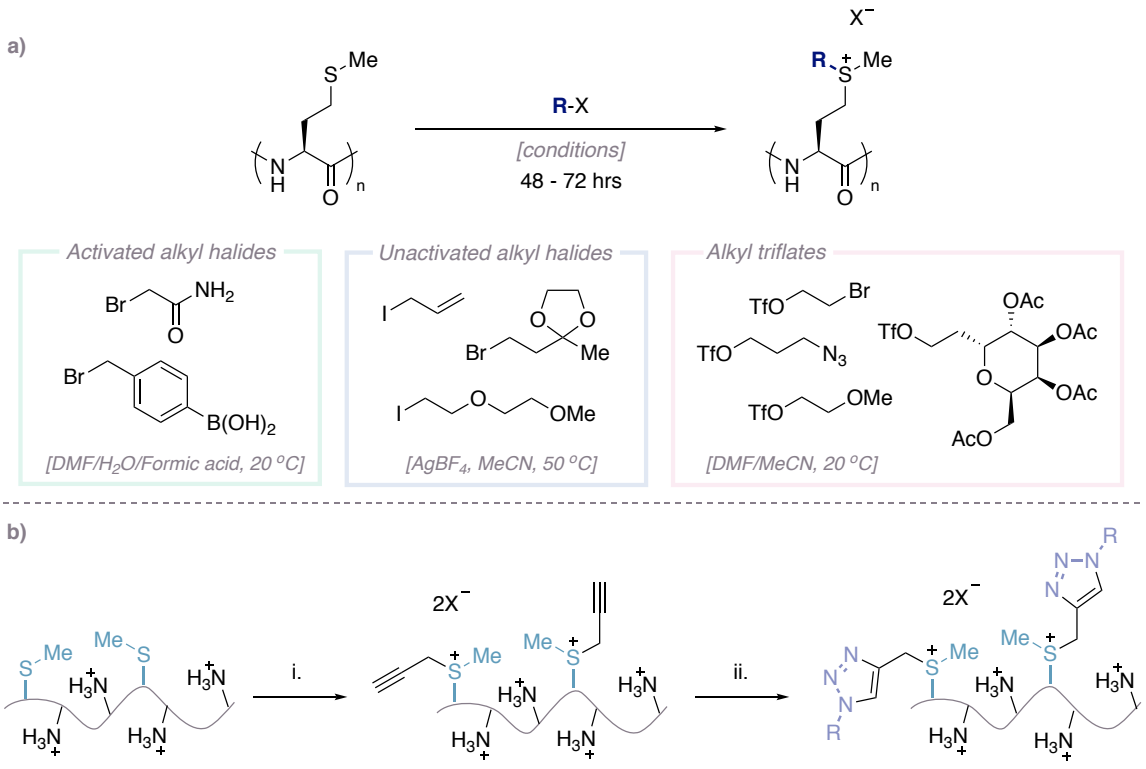
While a number of reports for the dealkylation of *S*-alkyl sulfonium residues have been demonstrated, showing the reversible nature of these transformations, conflicting reports have indicated that these groups are either inert or prone to decomposition under similar conditions.[173,183–185]



Scheme 24 Examples of the reversible alkylation of methionine

Deming and co-workers sought to expand the scope of methionine alkylation for the synthesis of polypeptides containing diverse and reactive functionality.[186] They aimed to develop a superior alkylative strategy using a polypeptide formed only of methionine residues, known as poly(Met). Despite the poor solubility of poly(Met) in many solvents, facile alkylation was achieved using a range of alkylating reagents, providing a high yielding process to form stable,

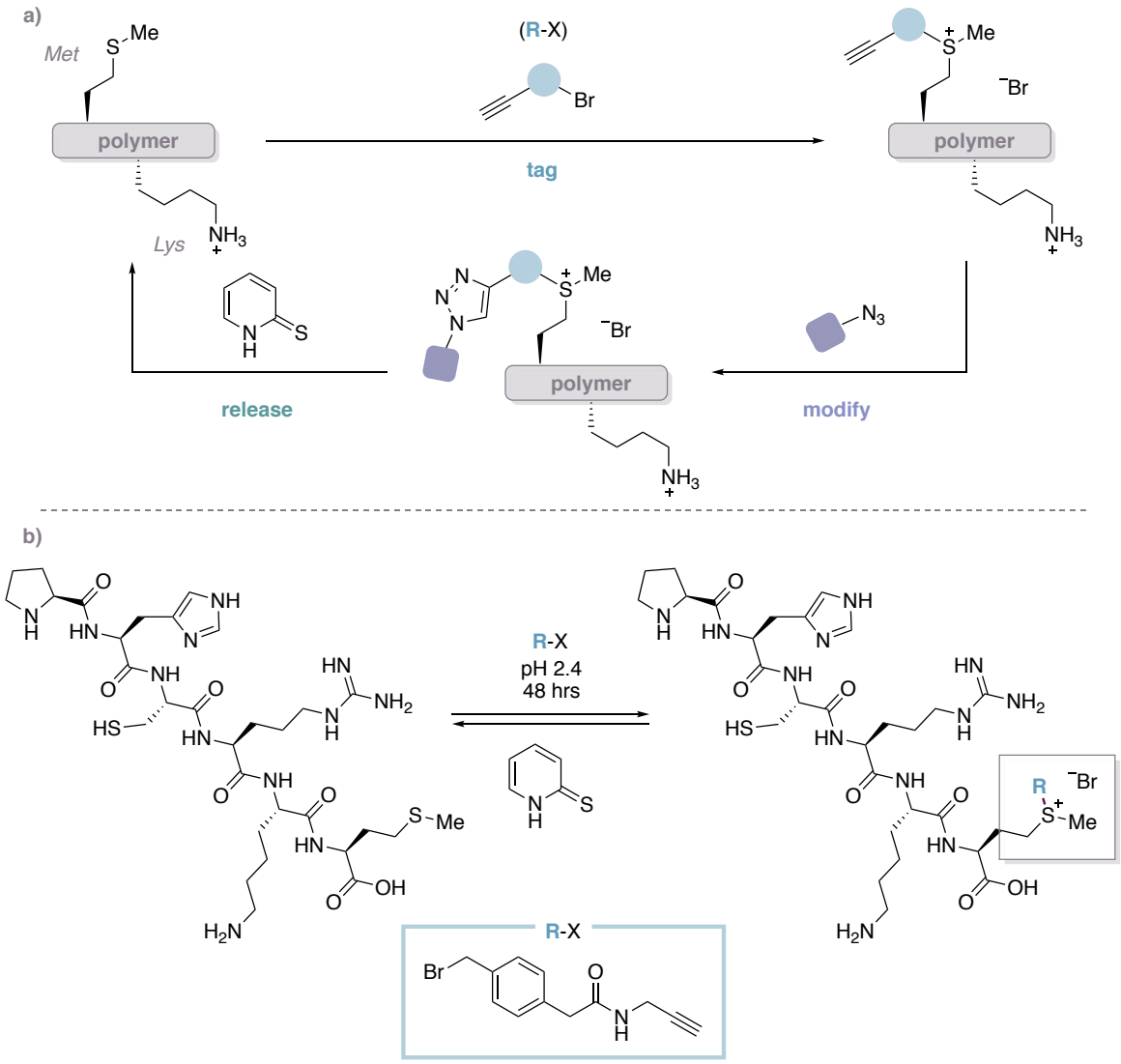
easily isolable products (Scheme 25a). Improving on previous strategies which generally employed only activated alkyl bromides and iodides, Deming demonstrated that, in a similar fashion to Luehr (Scheme 24c), the addition of silver tetrafluoroborate could promote the complete conversion of methionine alkylation with unactivated alkyl halides. Alternatively, to avoid the use of silver salts, functional alkyl triflates could be used in organic solvents for efficient alkylation of poly(Met). To exemplify the selectivity of the methionine alkylation procedure a polymer of methionine and lysine residues was synthesised and selectively alkylated at methionine. This was achieved through complete protonation of lysine residues under pH control (Scheme 25b). Interestingly, the alkyne groups introduced into the polymer could be quantitatively reacted with azide-terminated PEG chains in a subsequent bioorthogonal reaction. Importantly, this 'secondary click' strategy established the stability of these trialkyl sulfonium species and their compatibility with existing 'click' conjugation methods. Although the reaction displayed a broad scope of alkylating reagents, modification of methionine required the incorporation of organic solvents and long reaction times.



Scheme 25 a) Alkylation of poly(methionine) with a range of alkylating reagents b) Selective methionine functionalisation of synthetic copolymer; reagents i. propargyl bromide, 0.2M formic acid, pH 2.4 (94%) ii) α-methoxy-ω-azidoethyl-poly(ethyleneglycol), $CuSO_4$, ascorbic acid, PMDETA, H_2O (95%)

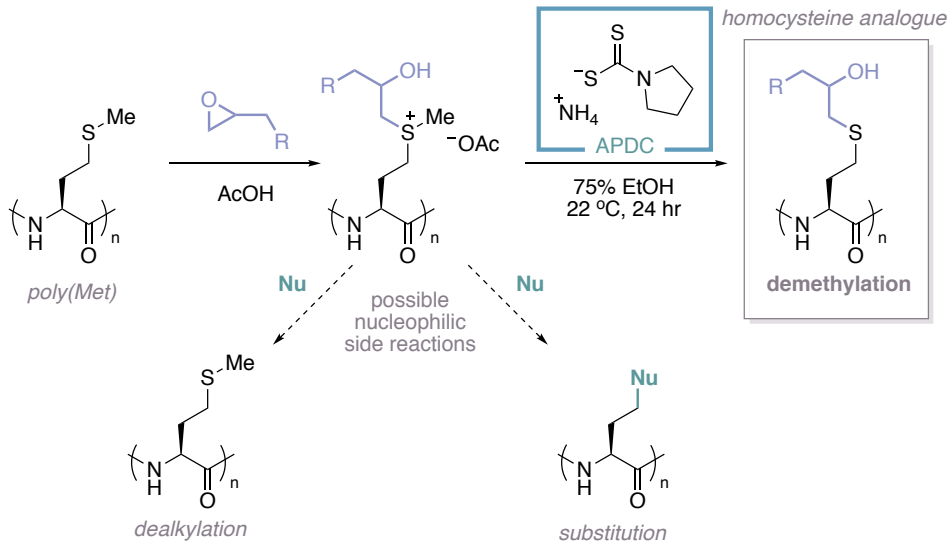
In further work, Deming described a more general platform for reversible, chemoselective alkylation of methionine using 2-mercaptopyridine **29** as the nucleophilic cleavage reagent.

This advancement proved distinct to previous methionine labelling strategies as it encompassed the general, triggered reversal of the alkylation. In contrast, earlier studies found that alkylating agents and nucleophilic cleavage reagents were substrate and functional group dependent, sometimes generating other products aside from regenerated methionine. Deming detailed an optimised alkylation strategy for the formation of stable sulfonium conjugates, enabling the introduction of bioorthogonal tags for subsequent modification and cleavage (Scheme 26a).[187] This was applied to a synthetic peptide which incorporated competing nucleophilic residues (Scheme 26b), demonstrating the selective nature of this reversible alkylation under pH control and alluding to its potential application in the controlled release of therapeutic peptides.



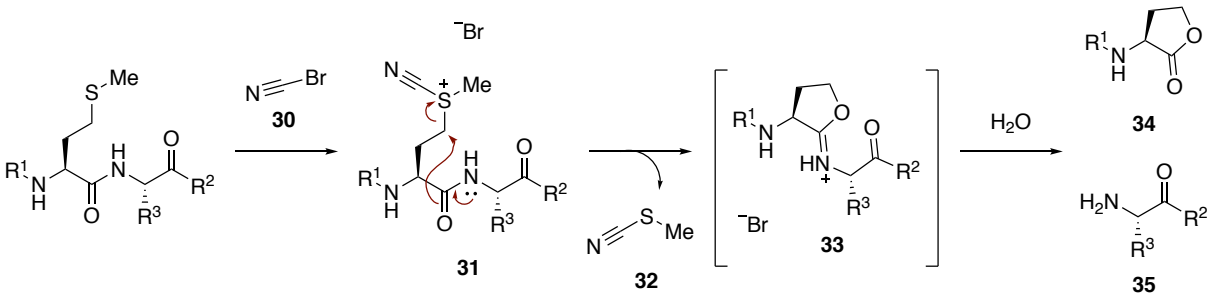
Scheme 26 a) Schematic for the alkylation, subsequent bioorthogonal modification and release of methionine b) Reversible alkylation of a short peptide containing highly nucleophilic, competing residues

Recently, Deming also revealed that alkylated methionine residues could be selectively demethylated using ammonium pyrrolidinedithiocarbamate (APDC) resulting in a neutral thioether with increased stability (Scheme 27).[188] In addition, through the use of substituted epoxides as alkylating reagents, demethylation could form homocysteine analogues. The process tolerates a number of different functional groups including azides, monosaccharides and zwitterionic groups, and can even selectively distinguish methyl from ethyl substituents. However, the reaction is highly dependent on the solvent, requiring less polar, ethanol-rich solvent mixtures to achieve good conversion over 24 hours.



Scheme 27 Selective demethylation of alkylated methionine residues with APDC (competing pathways of dealkylation and nucleophilic subsitition are shown)

In contrast to the alkylative strategies reported thus far, which form stable sulfonium products and aim to evade decomposition, a distinct methionine functionalisation strategy set out to form unstable sulfonium intermediates for the selective cleavage of methionyl peptide bonds. Products resulting from the alkylation of methionine with iodoacetamide require elevated temperatures for decomposition to the corresponding homoserine lactone. Gross and Witkop deduced that an alternative reagent would be required to form a sulfonium species which would be sufficiently unstable in order to undergo intramolecular displacement at room temperature under physiological pH in aqueous solution. They showed that the reagent cyanogen bromide (CNBr, **30**) could react efficiently with methionine residues under acidic conditions at room temperature to form an unstable cyanosulfonium salt **31**, which could undergo facile intramolecular cyclisation and loss of methyl thiocyanate **32** to form the corresponding iminolactone **33**. This unstable iminolactone is rapidly hydrolysed to two peptidic fragments; the homoserine lactone **34**, and the amine fragment **35** (Scheme 28).[189,190]



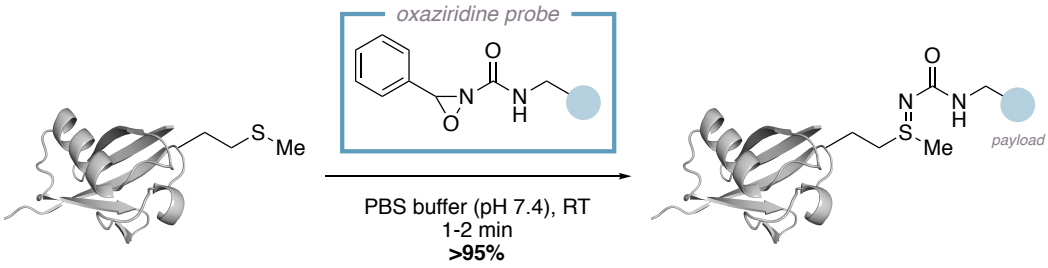
Scheme 28 Site-selective cleavage of methionyl peptide bonds using cyanogen bromide **30**

Although this degradative process is not truly a bioconjugation reaction, CNBr-mediated site-selective peptidic cleavage efficiently produces well-defined, smaller peptide fragments which can be analysed to give important insight into the composition of the larger initial protein scaffold. Cyanogen bromide cleavage of methionine-containing peptides and proteins has been invaluable in protein sequence determination and is used routinely in biological research.[191] The instability of the resulting cyanosulfonium salts demonstrate the importance of the high product stability which must be achieved in related bioconjugation reactions to avoid unwanted decomposition.

The strategies for selective modification of methionine residues reported thus far largely use conditions incompatible with biological systems. Methods often require elevated temperatures, proceed over long reaction times and form products with varying levels of stability. As such, many current techniques for methionine labelling represent chemical modification methods as opposed to true bioconjugation strategies. As an interesting target for bioconjugation, with low natural abundance and ancillary function, strategies for the selective functionalisation of methionine residues using biocompatible conditions remain highly sought after.

A recent advance was published in 2017 by Toste and Chang for the modification of methionine residues, representing a truly biocompatible strategy for methionine bioconjugation. The authors refer to the method as redox-activated chemical tagging (ReACT).[192] In contrast to the more traditional alkylative approaches, this oxidative sulfur imidation strategy uses oxaziridine reagents **36** to achieve highly selective, robust labelling in just 1-2 minutes, with a rate constant comparable to that of bioorthogonal reactions (18 $M^{-1}s^{-1}$)[8,192] (Scheme 29). This oxidative strategy exhibits high functional group tolerance in the presence of other nucleophilic residues, however cysteine residues are oxidised to the corresponding disulfide. The sulfimide products exhibit remarkable stability at varying pH, in the presence of disulfide reducing agents and also under bioorthogonal reaction conditions. Most notably, this strategy has proven selective for

methionine residues in both proteins and antibodies, allowing the facile synthesis of homogeneous antibody-drug conjugates (ADCs) and has also simplified proteomic profiling through identification of hyper-reactive methionine residues in whole proteomes.



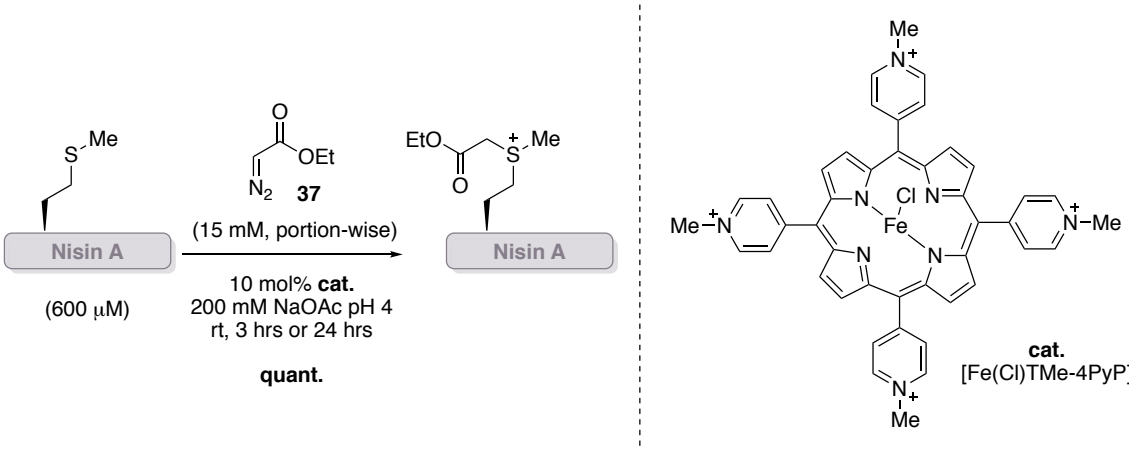
Scheme 29 Oxidative strategy for chemoselective methionine bioconjugation using oxaziridine reagents

1.4 Conclusion and outlook

Considering the pioneering work of Chang and Toste, it is evident that methionine is a valuable target for bioconjugation. Yet, while alkylative strategies remain the predominant method for chemical modification, inherent flaws such as long reaction times, stringent pH control and variable stability of sulfonium products preclude the applicability of current strategies to bioconjugation. A truly biocompatible, selective and efficient bioconjugation strategy is highly sought after, and many valuable techniques continue to be reported in the hope of achieving this goal.

The oxidative bioconjugation strategy published by Chang and Toste represents an excellent advance for methionine labelling. Although, with few strategies currently reported for methionine bioconjugation, there is chemical and biological need for further complementary techniques. Alkylative strategies forming cationic sulfonium moieties have been shown to be useful in the formation of cell-penetrating peptides as transporters for intracellular therapeutic delivery.[193] As such, the development of a biocompatible strategy of this nature would expand the repertoire of methods for bioconjugation and could prove invaluable in the advancement of biological applications.

A serendipitously discovered methionine alkylation strategy was reported concurrently to the publication of the work detailed in this thesis. As a novel, transition metal-catalysed methionine alkylation method, this strategy employed water soluble metalloporphyrins for selective bioconjugation under biocompatible conditions (Scheme 30).[194] While only reported on a single polypeptide system, this method provided a complementary strategy for methionine alkylation.



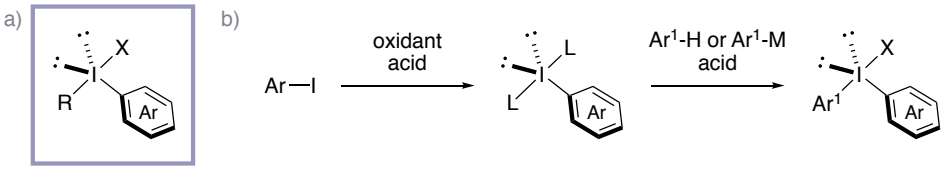
Scheme 30 Biocompatible methionine alkylation using metalloporphyrin catalysts for reaction with ethyldiazoacetate **37**

2 Development of a methionine-selective bioconjugation strategy

2.1 Hypervalent iodonium salts for bioconjugation

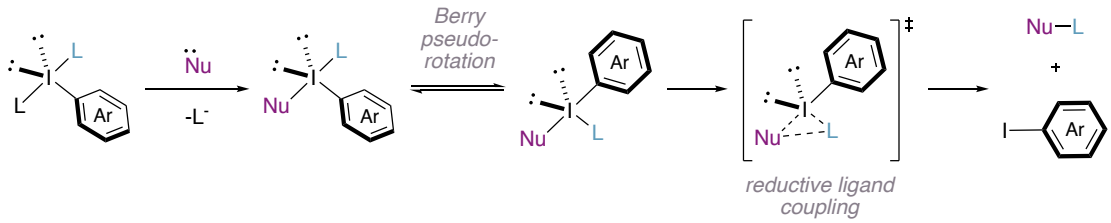
A number of strategies for bioconjugation have exploited hypervalent iodonium salts (λ^3-iodanes) for the selective and efficient introduction of exogenous functionality into peptides and proteins. Iodonium salts ($Ar(R)I^+X^-$) are a class of λ^3-iodane electrophiles in which two carbon based ligands (R, Ar) and a counter-anion (X) are bound to a central iodine(III) atom (Scheme 31a).[195] Despite being classed as 'salts', X-ray structural data for the majority of iodonium salts show secondary bonding between the iodine atom and the counter-anion (bond distance = 2.3-2.7 Å), resulting in a 'T-shaped' pseudo-trigonal bipyramidal geometry.[195] As such, the electrophilic iodine centre shares a hypervalent three-centre, four-electron bond with two axial ligands, which is longer and weaker than regular covalent bonds.[195] It is this 'hypervalent' bond that is responsible for the high electrophilic reactivity of λ^3-iodane species.

Synthesis of these reagents is usually achieved *via* initial oxidation of an aryl iodide to an iodine(III) species. Subsequent ligand exchange with an arene, organometallic reagent or other suitable nucleophile can occur in the presence of an acid to yield the λ^3-iodane (Scheme 31b). The anion (X) usually originates from the acid used in the ligand exchange, such as boron trifluoride diethyl etherate (X = BF_4) or trifluoromethanesulfonic acid (X = OTf).[196]



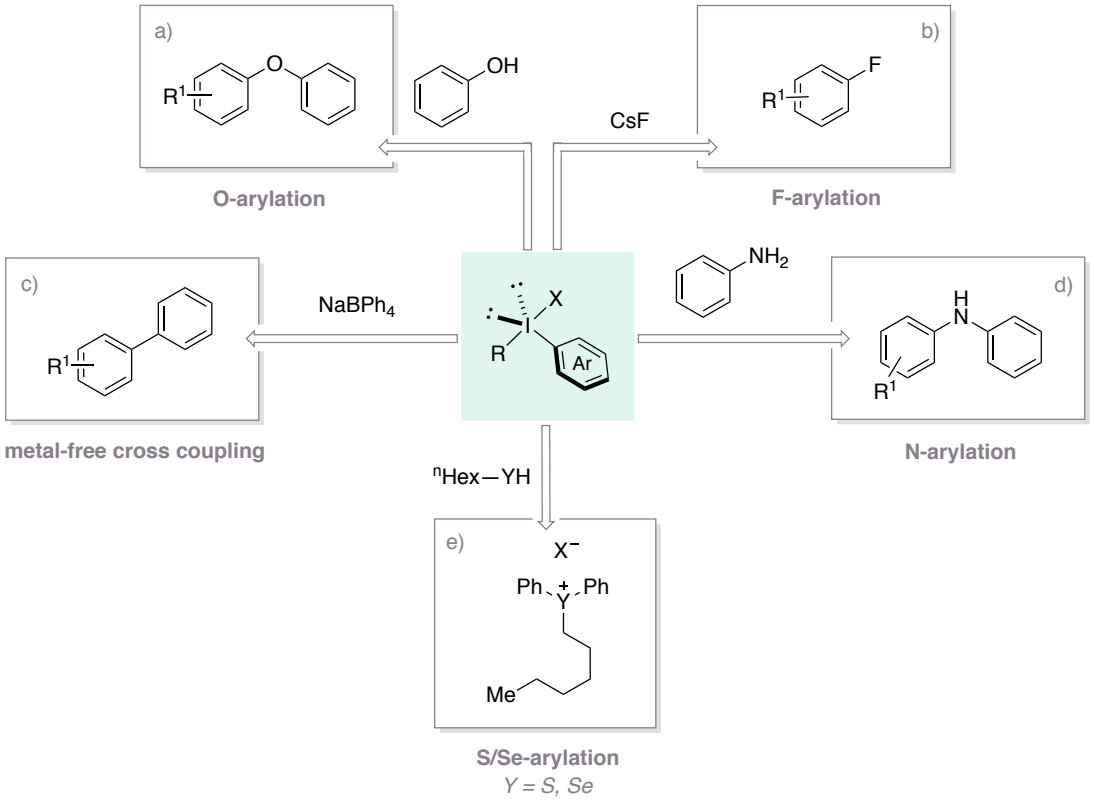
Scheme 31 a) General structure of hypervalent iodine reagents (R = carbon-based ligand, X = counter-anion). Considered 10-I-3 species under Martin-Arduengo designation [N-X-L] (where N = number of valence electrons on central atom, X = central atom, L = number of ligands on central atom)[197] b) General synthetic strategy for formation of diaryl iodonium salts

λ^3-iodanes react readily with nucleophiles to form a Nu–I bond upon displacement of a ligand. Reactions with unsymmetrical λ^3-iodanes give two T-shaped intermediates, which are in fast equilibrium through Berry pseudorotation.[198] The prevailing mechanism involves subsequent reductive ligand coupling of the nucleophile and the equatorial group to afford the product, Nu–L, releasing Ar–I (Scheme 32). Other mechanisms for the reactivity of hypervalent iodonium salts are possible and are discussed at length in a number of reviews.[195,196,199–201] The high reactivity of these λ^3-iodanes is explained by the exceptional leaving group ability of Ph–I.[202]



Scheme 32 General reactivity of λ^3-iodanes with nucleophiles

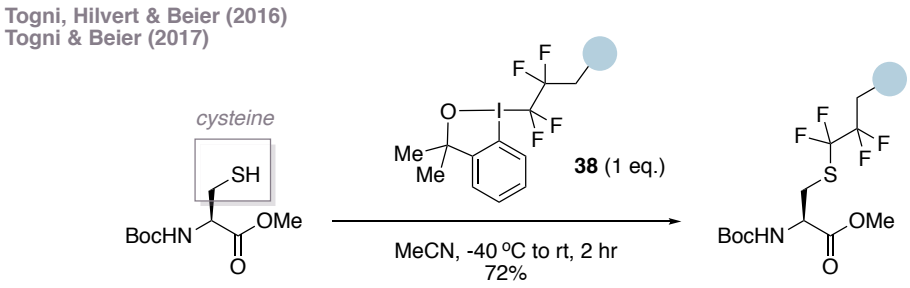
Iodonium salts are highly useful reagents for the electrophilic transfer of the ligand (R) to electron-rich organic substrates, and have been employed in the arylation of phenols (Scheme 33a),[203,204] fluoride anions (Scheme 33b)[205,206] sodium tetraphenylborate (Scheme 33c),[207–209] anilines (Scheme 33d),[210] and sulfur/selenium nucleophiles (Scheme 33e).[211–213] Diaryl iodonium salts, in which R = Ar, are one of the most common classes of λ^3-iodanes.



Scheme 33 Reactivity of iodonium salts with nucleophiles

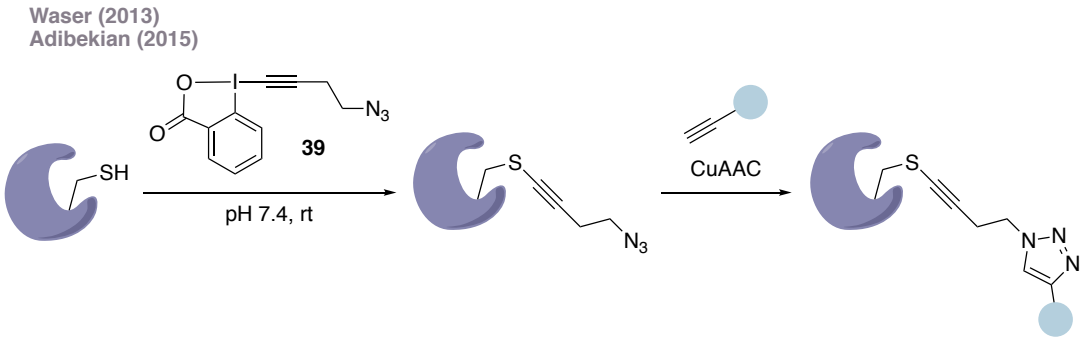
Most iodonium salts are air and moisture stable compounds which, depending on the nature of the counter-anion, possess some aqueous solubility.[196] They are generally mild and non-toxic and can react with certain nucleophiles with high selectivity. A number of protein functionalisation strategies using iodonium salts have been explored, utilising their desirable, biocompatible properties.

Togni, Hilvert and Beier were able to achieve the perfluoroalkylation of protected cysteine residues through the use of an iodine(III) heterocycle **38**, a five-membered ring variant of an λ^3-iodane reagent, in which iodine and oxygen are incorporated into the ring (Scheme 34).[214,215] The perfluoroalkylation proceeded at -40 °C in 2 hours, with the products shown to be stable and unreactive to external thiol nucleophiles. This strategy was applied to the labelling of exposed cysteine residues in a synthetic retro-aldolase enzyme.[215] The use of established maleimide based bioconjugation methods resulted in not only cysteine labelling, but also non-specific labelling of a lysine residue essential for enzymatic activity. In contrast, selective cysteine modification was achieved using **38**, leaving the lysine residue unmodified, highlighting the high selectivity which can be achieved using hypervalent iodine reagents.



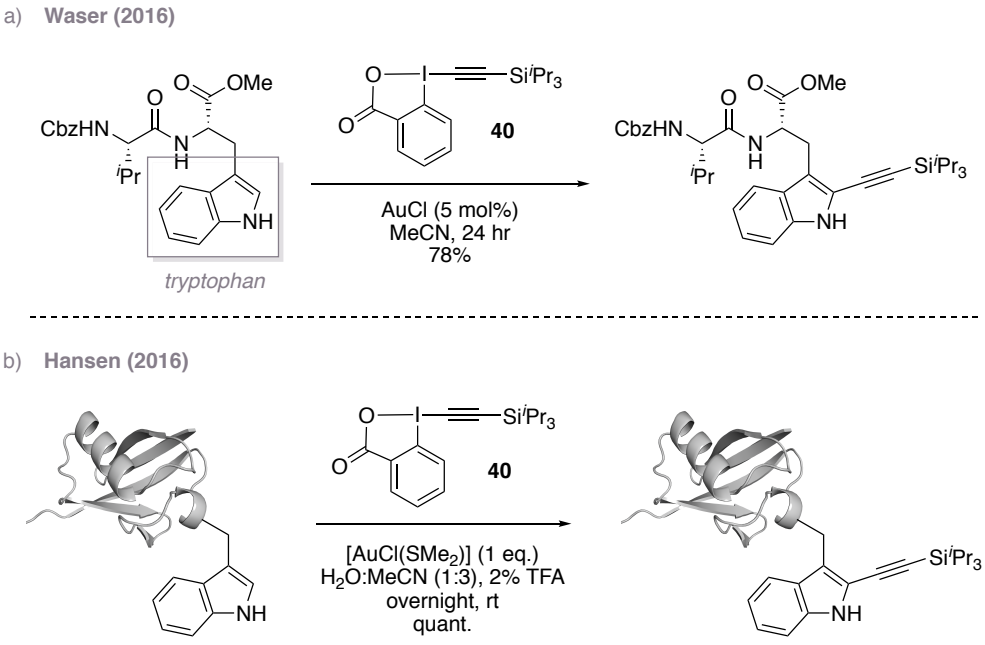
Scheme 34 Selective bioconjugation of cysteine residues using iodine (III) heterocycle

Waser and Adibekian also employed benziodoxole reagents for the alkynylation of cysteine residues (Scheme 35). Alkynylation was initially achieved on protected dipeptide substrates,[216] but could subsequently be applied to complex proteomes and even living cells.[217] The labelling was complete in just 30 seconds at room temperature and, through variation of the carbon-based ligand on the λ^3-iodane reagent **39**, could introduce a pendant azide moiety. A subsequent copper catalysed azide-alkyne cycloaddition reaction with an external alkyne reagent allowed the introduction of a functional payload to enable proteomic profiling in complex protein mixtures and in cells.



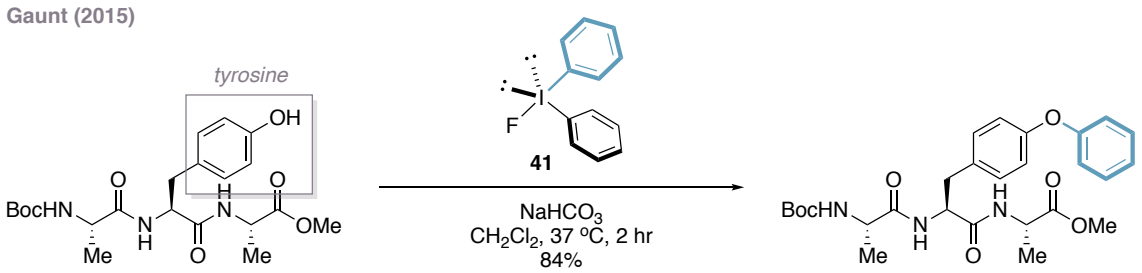
Scheme 35 Selective bioconjugation of cysteine residues within proteomes and living cells using iodine (III) heterocycle (CuAAC – Copper catalysed azide alkyne cycloaddition)

A similar benziodoxole reagent, **40**, was used for the alkynylation of the C2 position of tryptophan residues. Waser *et al.* described the gold-catalysed alkynylation of protected dipeptide substrates using **40** (Scheme 36a).[218] The reaction proceeded in good conversion over 24 hours and was selective for tryptophan over other aromatic amino acids. However, lower conversion was observed on peptides which were not protected at the C- and N-terminus. Hansen and co-workers detailed a similar gold-catalysed alkynylation strategy which could be applied to several polypeptides and proteins (Scheme 36b).[219] They employed the same benziodoxole reagent **40** for the modification of tryptophan at the C2 position, producing a single regioisomer which was confirmed using NMR and MS/MS analysis. No functionalisation was observed under strictly aqueous conditions, presenting a drawback of this strategy due to the requirement for organic solvent.



Scheme 36 Selective bioconjugation of tryptophan residues using iodine(III) heterocycles

Diaryl iodonium salts, whose use in the arylation of small molecules is well established, have also been used as reagents for peptide functionalisation. Gaunt and co-workers reported the use of diaryliodonium salt **41** for the O-arylation of a tyrosine residue in a protected tripeptide (Scheme 37).[204] Although the reaction proceeded in 2 hours, organic solvents and relatively high concentrations (> 1 mM) were required, factors which would likely prove incompatible with biological systems.



Scheme 37 Functionalisation of a tyrosine residue in a model tripeptide using a diaryliodonium salt

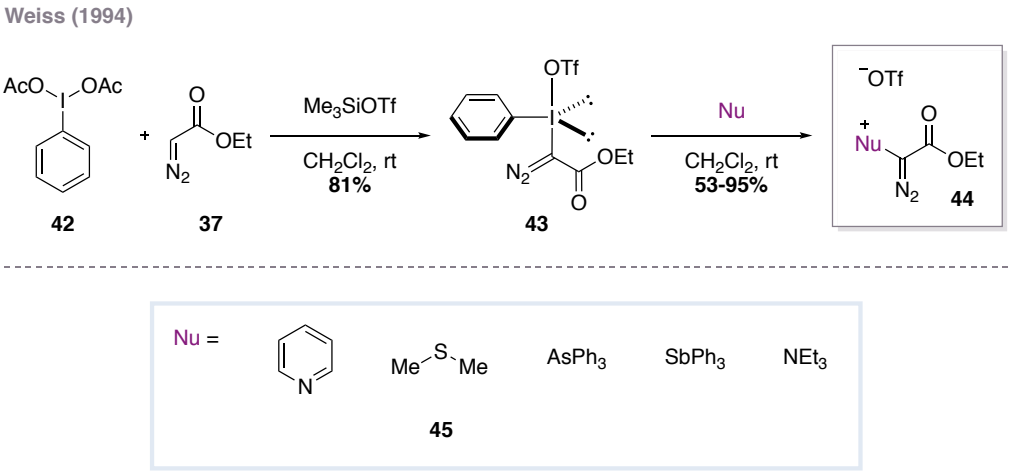
It has been shown that cyclic hypervalent iodine reagents can be valuable tools for selective bioconjugation, with a limited number of methods detailing the functionalisation of different amino acid residues. Due to the desirable properties of these λ^3-iodanes, including stability, partial water solubility and high selectivity, their application as reagents for bioconjugation has been explored. Thus far, there has been no report of the use of hypervalent iodine reagents for bioconjugation at methionine.

2.1.1 An α-aryliodonio diazo compound as a reagent for methionine bioconjugation

In 1994, Weiss *et al.* reported the synthesis and reactivity of a structurally remarkable iodonium salt **43** bearing an α-diazoacetate ligand (Scheme 38).[220] This was the first reported example of an α-aryliodonio diazo compound, and its synthesis was achieved in a single step from (diacetoxyiodo)benzene **42** and ethyldiazoacetate **37**. This novel iodonium salt displayed umpolung reactivity at the α-carbon of the diazoester motif due to the electrophilic nature of the iodine centre. As a result, **43** could react with a variety of neutral nucleophiles to afford stable α-substituted diazo products **44**.

Notably, iodonium salt **43** could react with dimethyl sulfide **45** resulting in the formation of a cationic sulfonium species. Due to the structural similarity of dimethyl sulfide to the methyl ethyl sulfide side chain of the amino acid methionine, our group proposed that the use of

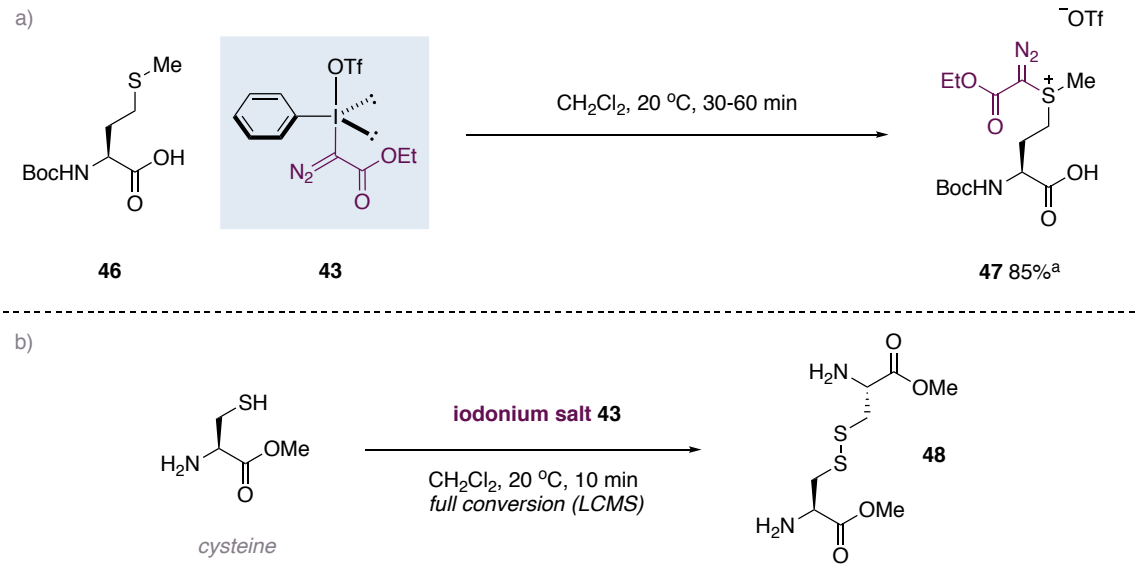
hypervalent iodine reagent **43** could present an opportunity for the development of a novel bioconjugation reaction at methionine. This would not only represent a distinct alkylative bioconjugation strategy, it would facilitate the introduction of reactive, non-native functionality into the biomolecule of interest. The newly introduced diazo moiety could also serve as a potential platform for further bioorthogonal transformations.



Scheme 38 Synthesis and reactivity of novel α-aryliodonio diazo compound **43**

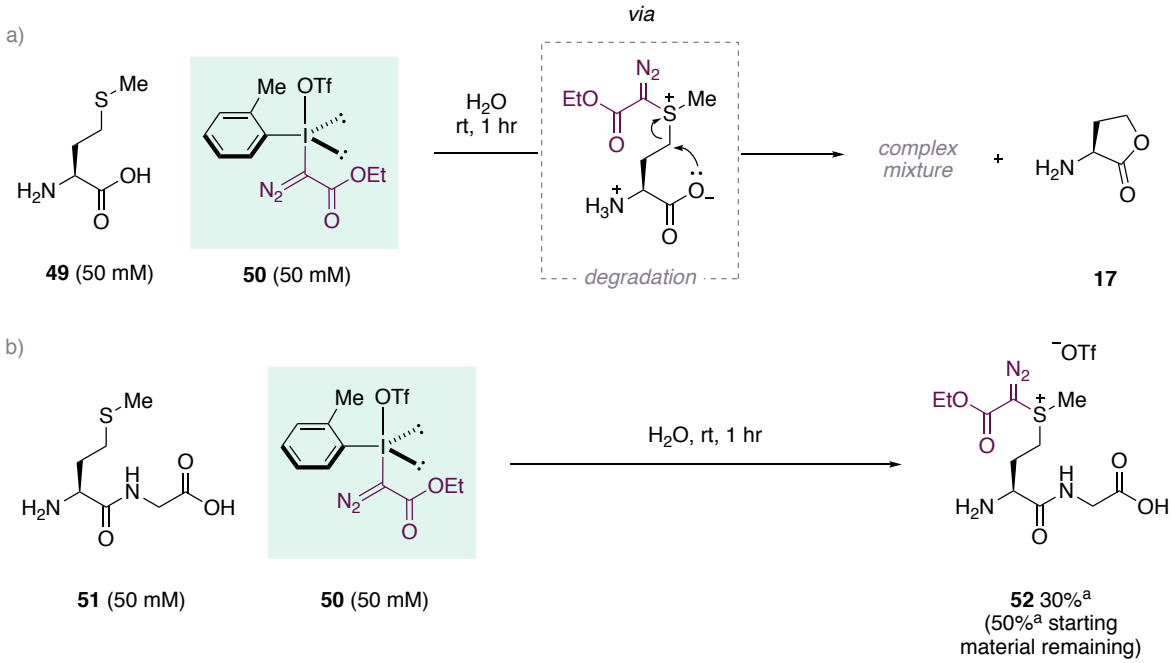
2.2 Previous work

Employing the original conditions reported by Weiss,[220] Dr M. G. Suero discovered that the reaction of a simple, protected methionine amino acid **46** with iodonium salt **43** afforded the sulfonium product **47** in 85% yield, determined by ^{1}H NMR (Scheme 39a). In order to gain an insight into the selectivity of the functionalisation, a control reaction on the amino acid cysteine was carried out (Scheme 39b). No sulfur alkylation occurred, with exclusive formation of the oxidised disulfide species **48** observed by LCMS analysis. Remarkably, the unprotected amine of the amino acid also remained untouched, demonstrating the selectivity of the iodonium salt for the sulfide of methionine over other competing nucleophiles.



Scheme 39 a) Initial reaction of methionine with Weiss iodonium salt **43** ([a] ^{1}H NMR yield); b) Exposure of cysteine to reaction conditions saw exclusive oxidation to disulfide **48**. Reactions carried out by Dr M. G. Suero

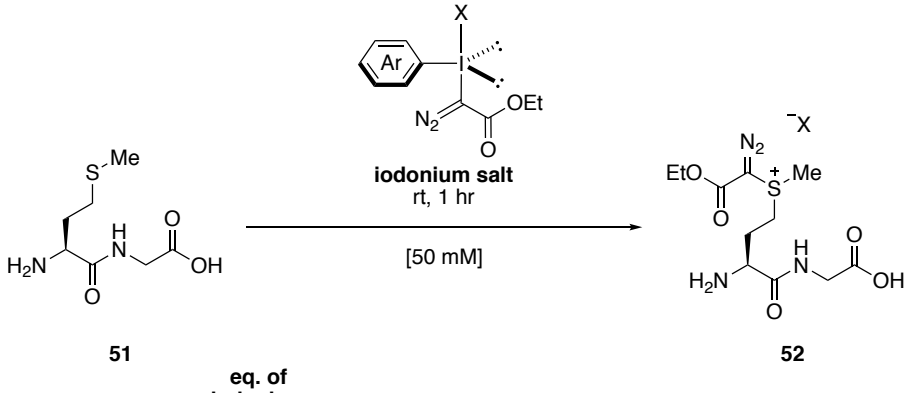
Dr M. G. Suero and Dr M. T. Taylor found that on switching the reaction solvent from CH_2Cl_2 to water, a complex mixture was observed for the reaction of methionine **46** with iodonium salt **43**. Early efforts towards the optimisation of the methionine functionalisation involved the alteration of the aryl group of the iodonium salt. Through the use of iodonium salt **50** with an *o*-tolyl aryl group, Dr M. T. Taylor found that a complex mixture was observed for the reaction with methionine **49** (Scheme 40a). Interestingly, homoserine lactone **17** could be isolated from the mixture, indicating that product formation had likely occurred but had proven unstable and undergone subsequent cyclisation with loss of a neutral thioether. A model dipeptide system (MetGly, **51**) proved more stable under the reaction conditions, forming the sulfonium product **52** in 30% yield (Scheme 40b). However, in this case the reaction was slow and did not go to completion, with 50% of the starting peptide remaining.



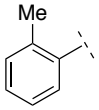
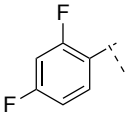
Scheme 40 Optimisation of methionine labelling in water; a) Degradation side-product **17** isolated *via* reverse phase HPLC b) Optimisation on model dipeptide substrate **51** ([a] ^{1}H NMR yield). Reactions carried out by Dr M. T. Taylor.

Further optimisation of the reaction identified that iodonium salts bearing an *o*-tolyl or 2,4-difluorophenyl moiety facilitated the highest yield of sulfonium product **52** (Table 1, Entries 2-5). Improvements to the yield could also be made through the addition of an aqueous solution of the iodonium salt as opposed to the solid reagent (Table 1, Entry 5). Various counterions were screened in the reaction (Ts_2N^-, TsO^-, OSO_3H^-, BF_4^-) and it was found that the salt with a BF_4^- counterion provided optimal results, increasing the yield to 95% (Table 1, entries 4-5) as well as improving the water solubility of the iodonium salt reagent. The optimal conditions (Entries 4 & 5, Table 1) afforded MetGly conjugate **52** in 95% yield over 1 hour. Due to the ease of purification and the stability of iodonium salt **54**, bearing a 2,4-difluorophenyl moiety and a BF_4 counterion, it was this reagent which was used by Dr M. T. Taylor for all further optimisation.

Table 1 Key results from the optimisation of methionine functionalisation through variation of the aryl group and counterion of the iodonium salt. Yields were calculated using ^{1}H NMR. Reactions carried out by Dr M. T. Taylor



51 **52**

Entry	Ar	X	eq. of iodonium salt	Solvent	[peptide]	yield[a]	comments
1		OTf	1	H_2O	50 mM	28%	
2		OTf	3	H_2O	50 mM	86%	
3		OTf	3	1:1 H_2O/MeCN	50 mM	90%	*iodonium salt added as solution*
4		BF_4	3	H_2O	50 mM	95%	*iodonium salt added as solution*
5		BF_4	3	H_2O	50 mM	95%	*iodonium salt added as solution*

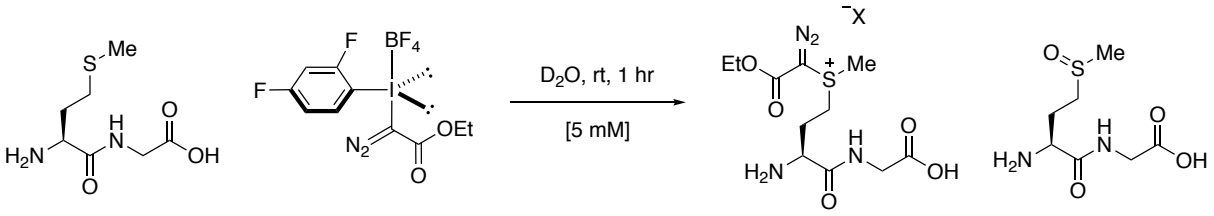
[a]Yield determined using ^{1}H NMR, with DMF as an internal standard.

In order for the reaction to be applicable to larger protein or enzymatic systems, it must operate successfully at low micromolar concentrations. Dr M. T. Taylor discovered that decreasing the concentration of peptide in the reaction (from 50 mM to 5 mM) led to an increase in an oxidised side-product, sulfoxide **55** (Table 2). This side-product was likely formed due to reaction of methionine with reactive oxygen species (ROS) produced from molecular oxygen such as the superoxide anion ($O_2^{-\bullet}$), hydroxyl radical ($^\bullet$OH) or hydrogen peroxide (H_2O_2).[162] To reduce such side reactions, Dr. M. T. Taylor screened a number of additives that were known to be ROS-scavengers (Table 2).[162,221,222]

While all of the ROS-scavengers tested reduced the amount of sulfoxide formed, the use of thiourea resulted in the highest yield, achieving >95% conversion to the desired sulfonium product **52**. Remarkably, it was observed that thiourea facilitated conversion to the desired product through both quenching of reactive oxygen species, and a rate acceleration effect. Additionally, the amine and carboxylic acid groups present within the dipeptide remained

untouched during the reaction, reinforcing the unreactive nature of the iodonium salt towards other nucleophiles.

Table 2 Screening of scavengers of ROS to reduce formation of oxidation side-product **55**. Yields calculated using ^{1}H NMR. Reactions carried out by Dr. M. T. Taylor



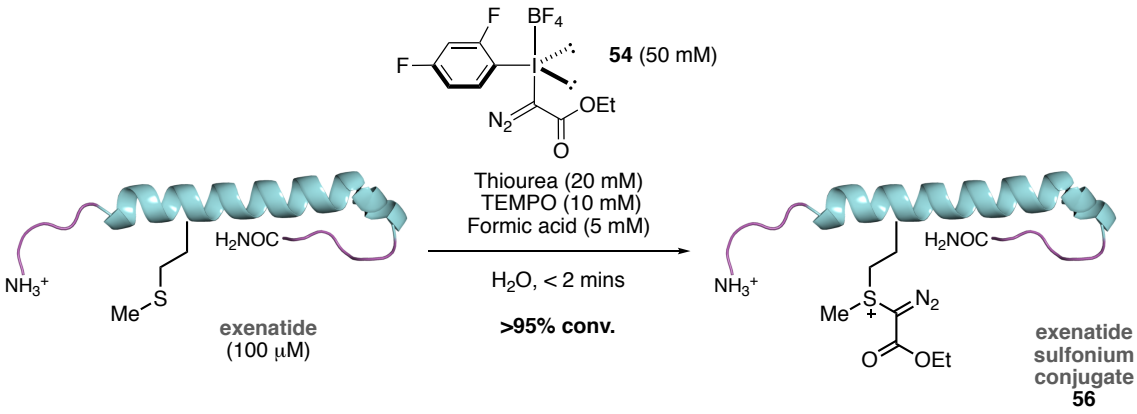
51 **54** **52** **55**

Entry	eq. of 54	Additive	yield 52 (yield 55)[a]	comments
1	10	—	47% (36%)	*12% **54** remaining*
2	10	TEMPO (10 mM)	70% (<5%)	*20% **54** remaining*
3	10	Ascorbic acid (10 mM)	73% (<5%)	*<5% **54** remaining*
4	10	Thiourea (10 mM)	>95% (<5%)	*<5% **54** remaining*

[a]Yield determined using ^{1}H NMR, using DMF as an internal standard.

Further optimisation was carried out by Dr M. T. Taylor in order to translate the methionine functionalisation from a small molecule system to a polypeptide, exenatide (Scheme 41). Exenatide (Byetta^{TM}) is a 39-amino acid synthetic peptide drug, of molecular weight 4186.6 Da, and is approved for the treatment of patients with type 2 diabetes mellitus.[223] It contains one methionine residue, alongside multiple different nucleophilic residues including lysine, serine and glutamic acid and has an α-helical secondary structural element. The selective functionalisation of the methionine residue of exenatide could be achieved in >95% conversion through treatment with **54** in combination with a selection of additives. As polypeptide architectures contain more readily oxidised functional groups than small molecule systems, an additional additive was required to prevent undesired oxidation of the polypeptide. Therefore, TEMPO was added to the reaction to act as a scavenger of ROS. Formic acid could aid in the prevention of any non-specific labelling through protonation of other nucleophilic residues such as lysine. Finally, thiourea was used and, although its role was not well-understood, appeared to accelerate the rate of the reaction dramatically. In the presence of thiourea the reaction reached completion in <2 minutes, compared to 30-60 minutes for the control reaction in the absence of thiourea.

Pleasingly, the reaction could be carried out at micromolar concentrations without competing levels of oxidation and could form a single product **56** in excellent conversion. The high selectivity achieved was significant considering the myriad other nucleophilic residues present in the polypeptide and indicated the exquisite selectivity of the iodonium salt **54** for the thioether group of methionine.



exenatide sequence

H_2N-His-Gly-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Leu-Ser-Lys-Gln-Met-Glu-Glu-Glu-Ala-Val-Arg-Leu-Phe-Ile-Glu-Trp-Leu-Lys-Asn-Gly-Gly-Pro-Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser-$CONH_2$

Scheme 41 Optimised conditions for the functionalisation of methionine within model polypeptide exenatide. Conversion determined by LCMS total ion count (TIC). Reaction carried out by Dr M. T. Taylor.

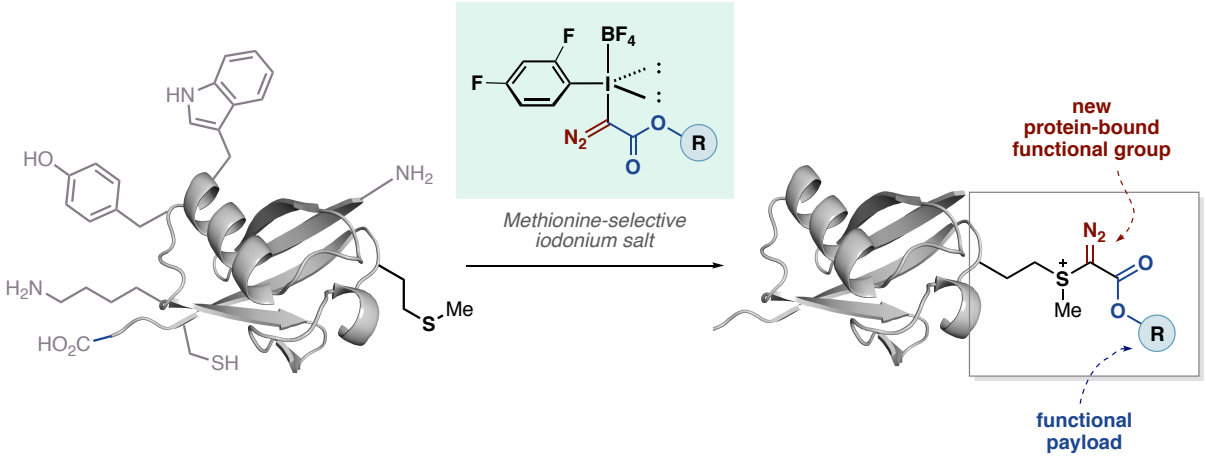
Overall, the initial results demonstrated the excellent selectivity of the reaction, implied fast kinetics and operated at low concentrations and under aqueous conditions. As a result, development of this methionine-selective strategy could provide a biocompatible method for the alkylative functionalisation of methionine residues within polypeptides and proteins.

2.3 Project aims

As discussed in Section 1.3, methionine presents an interesting target for bioconjugation due to its low natural abundance and ancillary protein function. Currently, there are no alkylative strategies for the bioconjugation of methionine which can operate under biocompatible conditions.

The goal of this project was to further optimise and develop a general platform for the alkylation of methionine using hypervalent iodine reagents. Building on the discovery and optimisation of this reaction on small molecule and model peptide systems, the aim was to expand the scope

of this reaction to a variety of polypeptide and protein substrates. Additionally, the aim was to synthesise a number of iodonium salts, varying the ester moiety (R), in order to transfer a range of functional payloads to the biomolecules of interest.



Scheme 42 Schematic representation of methionine-selective bioconjugation using hypervalent iodine reagents.

The introduction of a non-native diazo moiety using this technique also presented an opportunity for the development of subsequent bioorthogonal transformations at methionine residues. Hence, a final aim of this project was to explore the chemistry surrounding the β-sulfonium α-diazoester motif, aiming to exploit its high reactivity to introduce additional exogenous functionality. The goal was to combine bioconjugation reactions and bioorthogonal transformations for the development of a general methionine-selective protein functionalisation platform.

This project was carried out in collaboration with Dr M. T. Taylor, therefore due to the collaborative nature of this project, for clarity, all work undertaken is attributed by name in both the text and the schemes.

2.4 Results and discussion

2.4.1 Optimisation of methionine bioconjugation and isolation of sulfonium products

With initial optimised conditions in hand, attention was focused towards expanding the scope of polypeptide and protein substrates which could be efficiently and selectively labelled at methionine residues. While some polypeptide substrates showed promise for selective methionine labelling in mixed organic/aqueous solvent mixtures, successful reaction in fully aqueous conditions would be required to demonstrate a biocompatible approach.

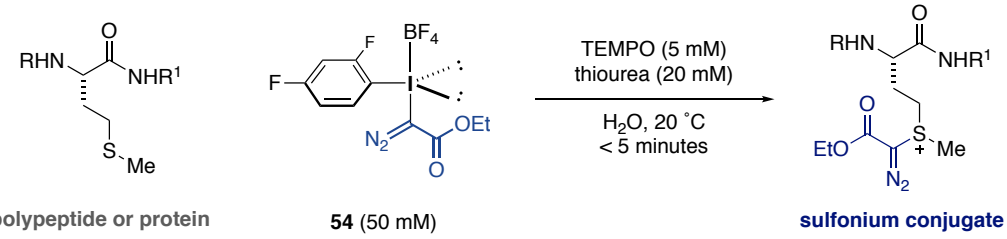
It was found that under completely aqueous conditions a number of polypeptide substrates could be functionalised in almost quantitative conversion (Scheme 43). These polypeptides represented small helical or unstructured polypeptides of molecular weight ranging from ~1.6-4.2 kDa. In most of these cases, the use of formic acid, previously used to reduce non-specific labelling, was not required. Avoiding the use of acid was advantageous as it provided conditions which would prove more biocompatible with protein systems.

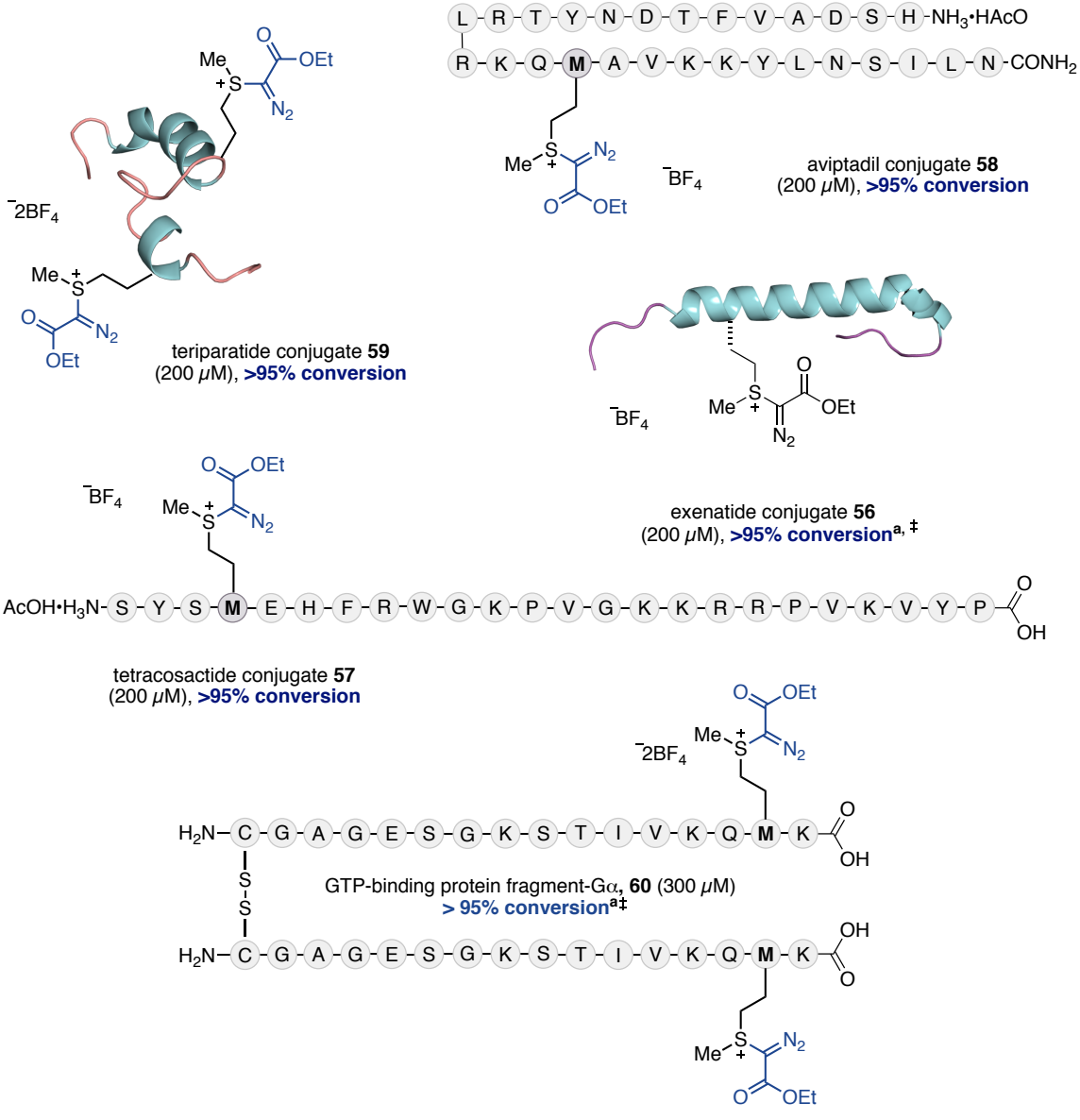
Aviptadil, a 28-amino acid synthetic peptide, exerts anti-inflammatory and immunomodulatory effects and is used in the treatment of pulmonary hypertension and erectile disfunction.[224] Selective labelling of aviptadil could be achieved despite the presence of several nucleophilic residues within its structure, including lysine, histidine and serine. Selectivity over the nucleophilic residue tyrosine, not present in exenatide, was also demonstrated on this substrate. Tetracosactide, a 24-amino acid synthetic peptide approved for medical use, was also tolerated well in the reaction.[225]

Teriparatide is the 34-amino acid, bioactive, N-terminal domain of parathyroid hormone and is used in the treatment of some forms of osteoporosis.[226] The sequence contains two methionine residues and the secondary structure includes two alpha helices. Interestingly, excellent conversion to the doubly-functionalised product **59** could be achieved with high selectivity. No singly-modified product was observed, confirming that all surface exposed methionine residues could be labelled using this technique.

The substrate GTP binding protein fragment Gα contains one methionine residue and one free cysteine residue. When subjected to the optimised reaction conditions by Dr M. T. Taylor, complete formation of the dimer was observed through disulfide bond formation at cysteine, with complete labelling of methionine residues. This represents an interesting example of

functionalisation in the presence of free cysteine residues, demonstrating complete selectivity for alkylation at methionine, with concomitant oxidation at cysteine.

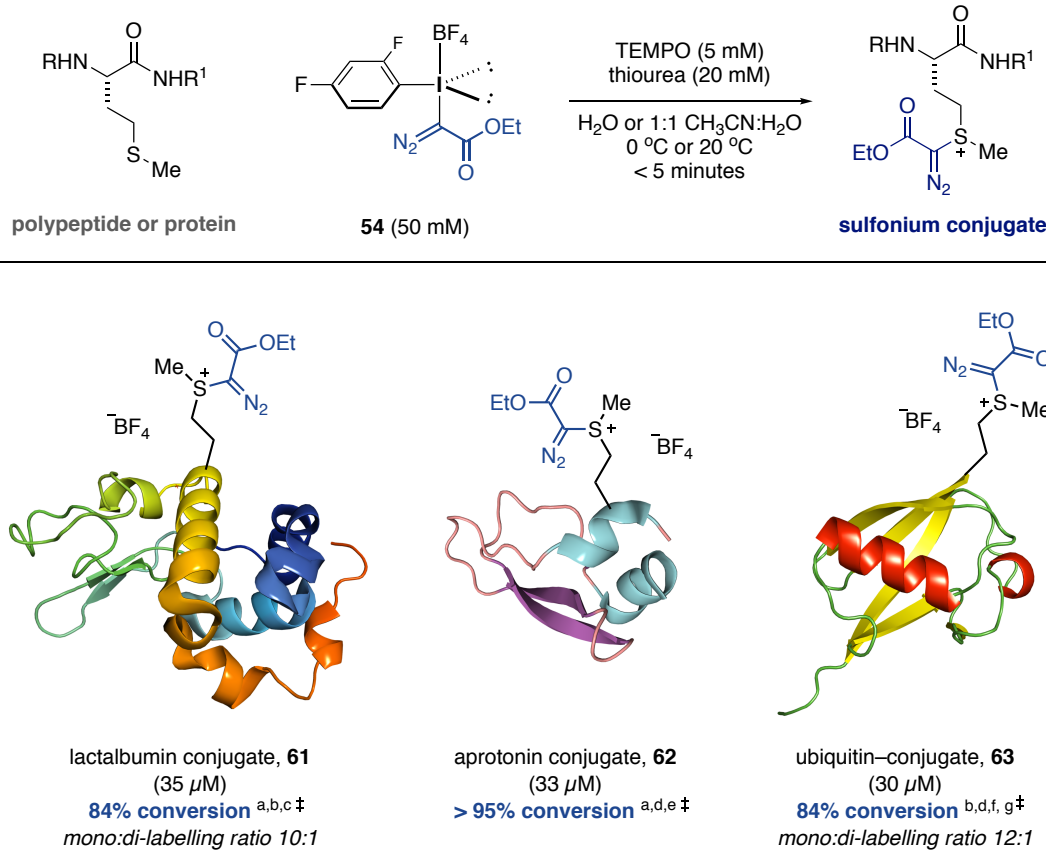




Scheme 43 Substrate scope of polypeptides for methionine-selective bioconjugation. Conversion determined *via* LCMS total ion count. [a] 5 mM Formic acid. [‡] Reaction conducted by Dr M. T. Taylor.

Attention was next focused on the optimisation of protein substrates in order to demonstrate further applicability of this methodology. Dr M. T. Taylor was able to establish that methionine could be functionalised selectively in three protein systems of molecular weight ranging from ~6.6-14.2 kDa (Scheme 44). However, reactions for the formation of lactalbumin conjugate **61** and ubiquitin conjugate **63** required mixed organic/aqueous solvent systems of 1:1 CH_3CN:H_2O and achieved a lower conversion to product of 84%. The labelling of these substrates in 100% H_2O was attempted, varying reaction time and temperature, although unfortunately delivered significantly lower conversions to the desired products. Nevertheless, successful functionalisation of these protein substrates demonstrated the increased complexity which could be tolerated by the methionine bioconjugation reaction, including secondary structural elements such as β-strands and β-sheets. Additionally, both aprotinin and lactalbumin contain multiple disulfide bonds which were unreactive to the iodonium salt **54**, demonstrating the high selectivity of this reagent for methionine.

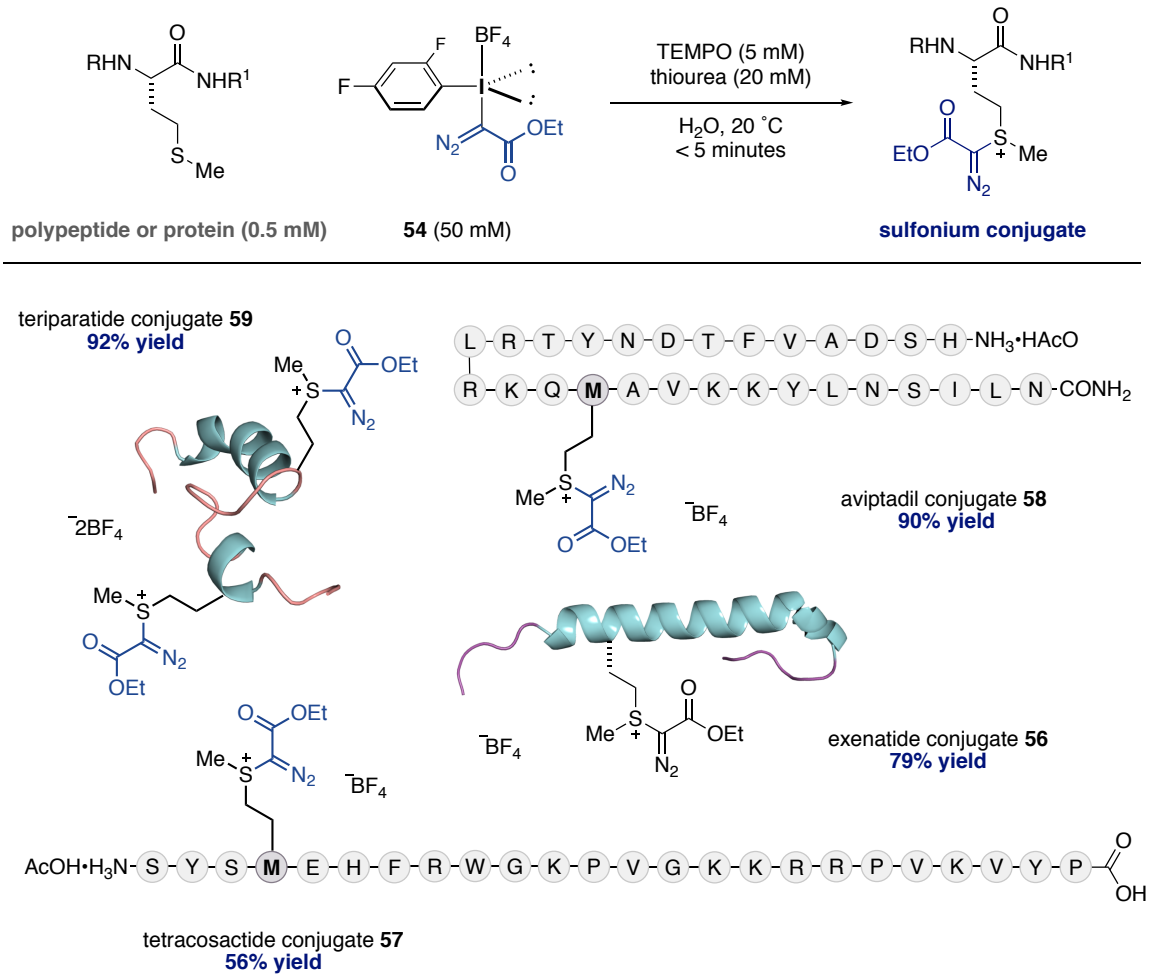
In the case of protein RNA-ase B, in which all methionine residues are buried within the folded protein structure, bioconjugation was unsuccessful. This lack of reactivity highlighted the inherent selectivity of the process for exposed methionine residues. The selectivity for exposed residues was also evident in the case of ubiquitin, in which the methionine residue is partially buried. The limited solvent exposure caused the rate of functionalisation to be retarded to the point that oxidation to the methionine sulfoxide became competitive. As a result, the labelling required deoxygenated conditions to achieve efficient conversion to sulfonium conjugate **63**.



Scheme 44 Substrate scope of proteins for methionine-selective bioconjugation. Conversion determined *via* LCMS total ion count. ‡ Reactions conducted by Dr M. T. Taylor. [a] 0 °C. [b] 1:1 CH_3CN:H_2O. [c] 10 mM HCl. [d] 5 mM Formic acid. [e] H_2O. [f] 20 °C. [g] Reaction carried out under N_2, using N_2-degassed solvent.

2.4.2 Structural analysis of β-sulfonium α-diazoester conjugates

It was envisaged that the selectivity of the methionine labelling could be probed through characterisation using tandem mass spectrometry (MS/MS). Purification of polypeptide sulfonium conjugates was required in order to carry out MS/MS analysis. Therefore, polypeptide-sulfonium conjugates **56-59** were isolated on a larger scale (~ 2 mg) using reverse phase semi-preparative HPLC. In the case of these substrates, the reaction proceeded in almost quantitative conversion. Therefore, isolation required only the separation of the sulfonium conjugate from by-products resulting from the consumption of the iodonium salt reagent and the additives such as thiourea. Yields of isolated products were determined by UV concentration analysis using A_{280} absorbance of chromophoric residues tyrosine and tryptophan. Polypeptides **56-59** were isolated in yields ranging from 56-92% (Scheme 45). Isolation of these functionalised peptide scaffolds in synthetically useful yields demonstrates the stability of the sulfonium products to modification on a larger scale and subsequent purification *via* HPLC.



Scheme 45 Isolated yields for larger scale reactions with polypeptides **56-59**. Yield determined by concentration analysis using A_{280} absorbance.

MS/MS is a technique frequently used for sequence determination of peptides and proteins and for the identification of post-translational modifications.[227] Charge-directed cleavages of peptidic bonds generate ions, described using Roepstorff nomenclature, which can be used to assign the sequence of a peptide or protein. Fragment ions retaining the positive charge on the N-terminus are termed a-, b- or c-type, while ions retaining the positive charge on the C-terminus are termed x-, y- or z-type. These ions occur with differing probabilities, with b- and/or y-type ions the most commonly observed (Figure 3).[228]

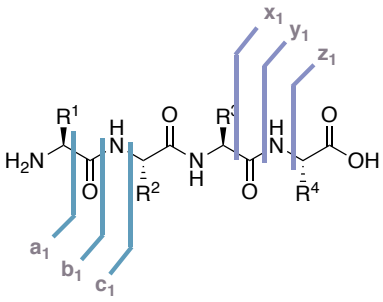
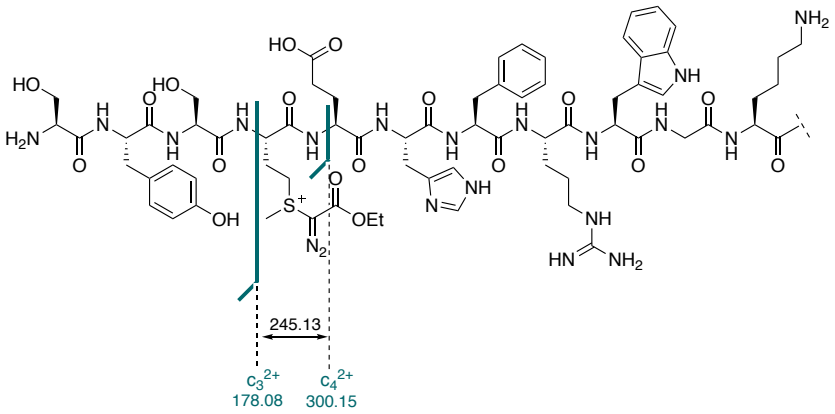


Figure 3 Roepstorff nomenclature for the fragmentation ions formed through backbone cleavage of protonated peptides. Fragment ions retaining the positive charge on the N-terminus are termed a-, b- or c-type, ions retaining the positive charge on the C-terminus are termed x-, y- or z-type.[228]

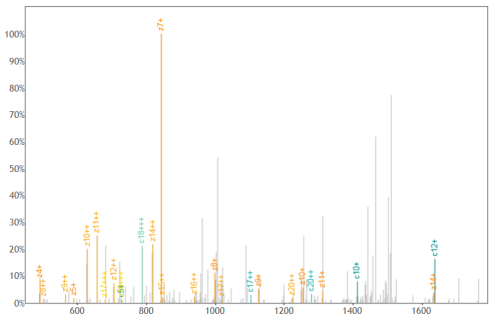
Purified sulfonium conjugates **56-59** were subjected to MS/MS analysis, carried out by Dr H. J. Lewis and Dr M. J. Edgeworth (AstraZeneca/MedImmune). Due to the relatively small size of the polypeptides, MS/MS analysis could be carried out on the intact modified polypeptide and did not require tryptic digestion. The MS/MS data for tetracosactide conjugate **57** is presented in Scheme 46.

Diagnostic fragment ions c_3^{2+} and c_4^{2+} were observed, the difference of which was equal to the sum of the methionine residue plus the desired modification (Scheme 46). Therefore, methionine was confirmed as the site of labelling. In the case of all polypeptides **56-59**, MS/MS data concluded that modification had taken place exclusively at methionine. Additional MS/MS data for polypeptide conjugates **58** and **59** can be found in Section 5.3.1.

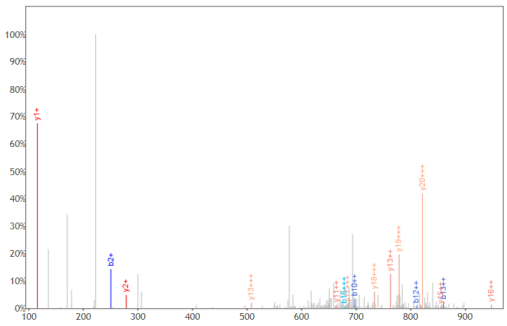


Sequence: SYS**M**EHFRWGKPVGKKRRPVKVYP

b+	b2+	b3+	c+	c2+	c3+	#	Seq	#	y+	y2+	y3+	z+	z2+	z3+
88.0393	44.5233	30.0180	105.0659	53.0366	35.6935	1	S	24						
251.1026	126.0550	84.3724	268.1292	134.5682	90.0479	2	Y	23	2958.6508	1479.8290	986.8884	2942.6321	1471.8197	981.5489
338.1347	169.5710	113.3831	355.1612	178.0842	119.0586	3	S	22	2795.5875	1398.2974	932.5340	2779.5687	1390.2880	927.1944
582.2700	291.6387	194.7615	599.2966	300.1519	200.4370	4	M	21	2708.5554	1354.7814	903.5233	2692.5367	1346.7720	898.1838
711.3126	356.1600	237.7757	728.3392	364.6732	243.4512	5	E	20	2464.4200	1232.7137	822.1449	2448.4013	1224.7043	816.8053
848.3716	424.6894	283.4620	865.3981	433.2027	289.1376	6	H	19	2335.3775	1168.1924	779.1307	2319.3587	1160.1830	773.7911
995.4400	498.2236	332.4848	1012.4665	506.7369	338.1604	7	F	18	2198.3185	1099.6629	733.4444	2182.2998	1091.6535	728.1048
1151.5411	576.2742	384.5185	1168.5676	584.7875	390.1941	8	R	17	2051.2501	1026.1287	684.4216	2035.2314	1018.1193	679.0820
1337.6204	669.3138	446.5450	1354.6469	677.8271	452.2205	9	W	16	1895.1490	948.0781	632.3879	1879.1303	940.0688	627.0483
1394.6419	697.8246	465.5521	1411.6684	706.3378	471.2277	10	G	15	1709.0697	855.0385	570.3614	1693.0510	847.0291	565.0218
1522.7368	761.8720	508.2505	1539.7634	770.3853	513.9260	11	K	14	1652.0482	826.5278	551.3543	1636.0295	818.5184	546.0147
1619.7896	810.3984	540.6014	1636.8161	818.9117	546.2769	12	P	13	1523.9533	762.4803	508.6559	1507.9346	754.4709	503.3164
1718.8580	859.9326	573.6242	1735.8845	868.4459	579.2997	13	V	12	1426.9005	713.9539	476.3050	1410.8818	705.9445	470.9654
1775.8795	888.4434	592.6313	1792.9060	896.9566	598.3069	14	G	11	1327.8321	664.4197	443.2822	1311.8134	656.4103	437.9426
1903.9744	952.4908	635.3297	1921.0010	961.0041	641.0052	15	K	10	1270.8106	635.9090	424.2751	1254.7919	627.8996	418.9355
2032.0694	1016.5383	678.0280	2049.0959	1025.0516	683.7035	16	K	9	1142.7157	571.8615	381.5767	1126.6969	563.8521	376.2372
2188.1705	1094.5889	730.0617	2205.1970	1103.1022	735.7372	17	R	8	1014.6207	507.8140	338.8784	998.6020	499.8046	333.5388
2344.2716	1172.6394	782.0954	2361.2982	1181.1527	787.7709	18	R	7	858.5196	429.7634	286.8447	842.5009	421.7541	281.5051
2441.3244	1221.1658	814.4463	2458.3509	1229.6791	820.1218	19	P	6	702.4185	351.7129	234.8110	686.3998	343.7035	229.4714
2540.3928	1270.7000	847.4691	2557.4193	1279.2133	853.1446	20	V	5	605.3657	303.1865	202.4601	589.3470	295.1771	197.1205
2668.4877	1334.7475	890.1674	2685.5143	1343.2608	895.8429	21	K	4	506.2973	253.6523	169.4373	490.2786	245.6429	164.0977
2767.5562	1384.2817	923.1902	2784.5827	1392.7950	928.8658	22	V	3	378.2023	189.6048	126.7390	362.1836	181.5954	121.3994
2930.6195	1465.8134	977.5447	2947.6460	1474.3267	983.2202	23	Y	2	279.1339	140.0706	93.7162	263.1152	132.0612	88.3766
						24	P	1	116.0706	58.5389	39.3617	100.0519	50.5296	34.0221



ETD fragmentation spectrum c and z ions



HCD fragmentation spectrum b and y ions

Scheme 46 MSMS analysis of tetracosactide conjugate **57**. ETD and HCD fragmentation spectrum show the spectra resulting from two different ionisation techniques. The highlighted boxes in the table and the corresponding highlighted peaks in the ETD/HCD spectra indicate a match of the theoretical values to the observed fragment ions for **57**.

The secondary structure of peptides and proteins is often crucial to their biological function and activity. Therefore, circular dichroism (CD) spectra were obtained for purified polypeptide conjugates **56-59** to determine whether the secondary structure and folding had been perturbed by the modification. For aviptadil and tetracosactide conjugates **58** and **57**, the secondary structure was less well-defined and changes in CD spectra were less apparent (see Section 5.3.1). However, the retention of secondary structure was evident in the case of teriparatide conjugate **59** (Figure 4). Clear overlap of the spectra resulting from the modified and unmodified polypeptides provided evidence that the secondary structure was not significantly affected by modification at methionine residues. Pleasingly, negative bands could be observed for the two α-helices in the characteristic regions of 222 nm and 208 nm,[229] confirming that the distinctive structure had been retained. This highlights the suitability of methionine as a target for bioconjugation.

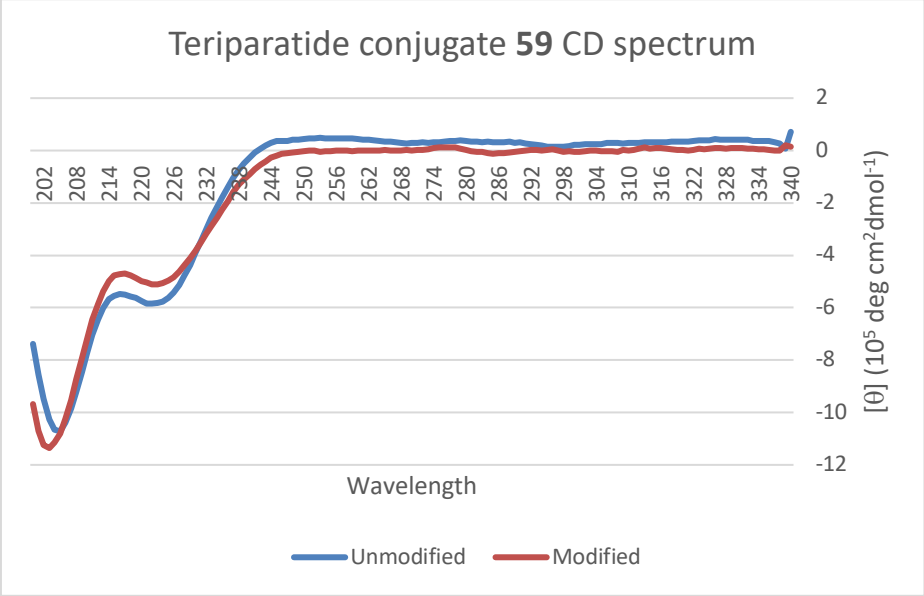


Figure 4 Circular dichroism spectra of teriparatide and modified teriparatide **59** in 5 mM pH 7.4 buffer

Isolation of the protein sulfonium conjugate of ubiquitin **63** was attempted. However, in contrast to the isolation of the polypeptide substrates, reactions on a larger scale, did not go to completion, with significant quantities of starting protein ubiquitin remaining. Nevertheless, isolation of the crude reaction mixture was pursued. The small difference in mass between the modified and the unmodified protein was reflected in poor separation between species during HPLC purification. As such, separation proved extremely difficult. A cationic sulfonium moiety is formed through this alkylative strategy, hence it was envisaged that separation of modified and unmodified species could be carried out using ion exchange chromatography, a

technique which separates compounds based on their differing affinity to an ionic stationary phase. With the assistance of Dr N. Huguenin, Fast Protein Liquid Chromatography (FPLC) purification was carried out on the crude reaction mixture of ubiquitin conjugate **63**. Isolation of pure material was achieved, with good separation from the ubiquitin protein using an ion exchange (IEX) column for FPLC (Section 5.3.1, Figure 39). As sodium chloride and ammonium acetate buffers were used for product elution, salts remained present in the isolated product sample and prevented CD analysis. Desalting of the sample was attempted but resulted in degradation of the conjugate.

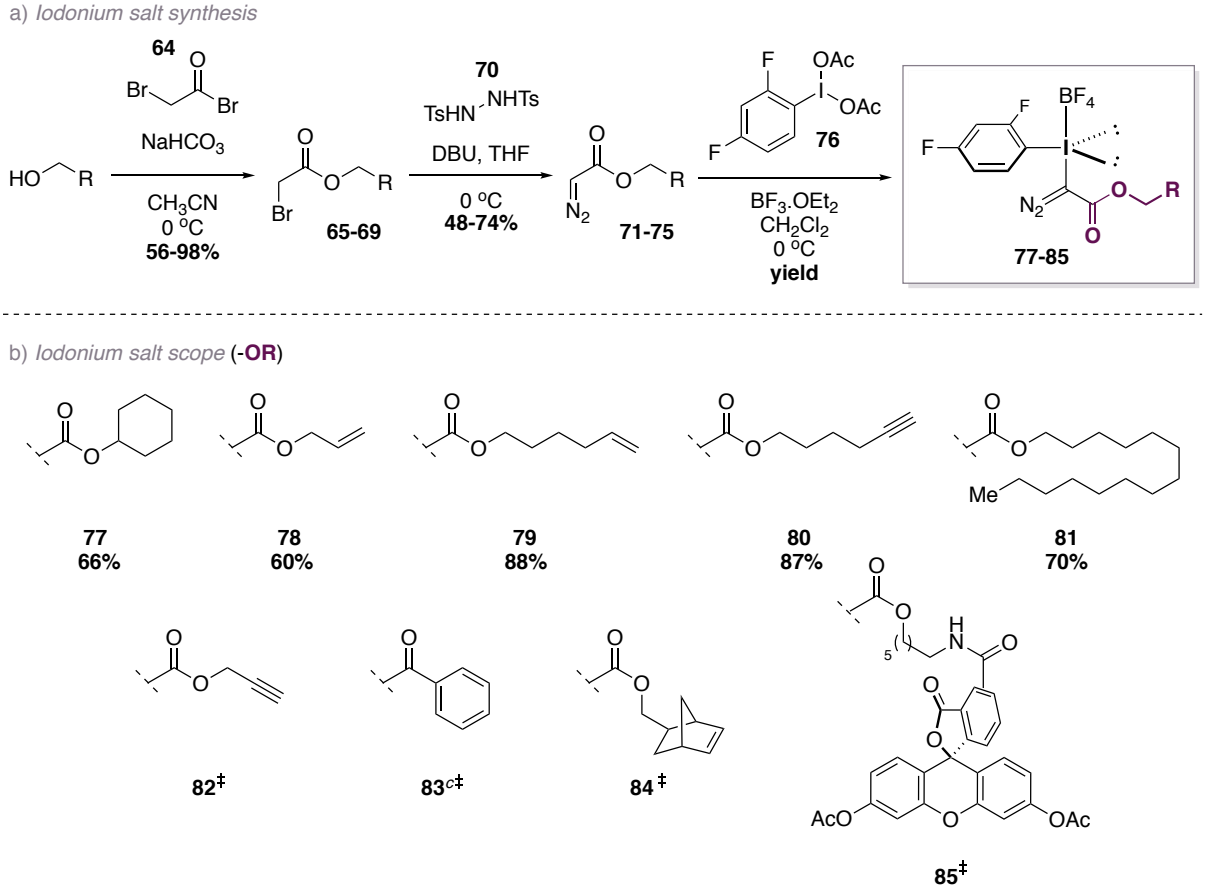
In contrast to the polypeptide sulfonium conjugates, MS/MS analysis could not be obtained for the larger protein substrates **61**-**63**, despite continued efforts. Nonetheless, selectivity could be inferred from the previous polypeptide examples.

2.4.3 Scope of iodonium salt transferring reagents

Following the successful bioconjugation of several polypeptides and proteins, efforts focused on varying the functional payload which could be transferred to the biomolecule of interest. It was hypothesised that this could be achieved through variation of the ester group of the iodonium salt reagent. Starting from commercially available alcohols, reaction with bromoacetyl bromide **64** afforded the corresponding α-bromoacetates **65-69** in 56-98% yield (Scheme 47a). Reaction of **65-69** with *N,N'*-ditosylhydrazine **70** provided various α-diazoacetates in 48-74% yield. Subsequent reaction with (2,4-difluorophenyl)-λ^3-iodanediyl acetate **76**, in the presence of a Lewis acid furnished the desired iodonium salts **77**-**85** in 60-87% yield.

Iodonium salts were synthesised incorporating a number of different functional groups including alkenes (**78**, **79**), alkynes (**80**, **82**) or norbornene moieties (**84**) which provided the opportunity for subsequent bioorthogonal transformations[8] (Scheme 47b). An iodonium salt was synthesised incorporating a myristyl chain (**81**) which, when introduced into proteins, could mimic a post-translational modification.[4] Additionally, a fluorescein-derived iodonium salt (**85**) was synthesised by Dr M. T. Taylor which could potentially be used as a valuable reagent in cellular imaging.[5]

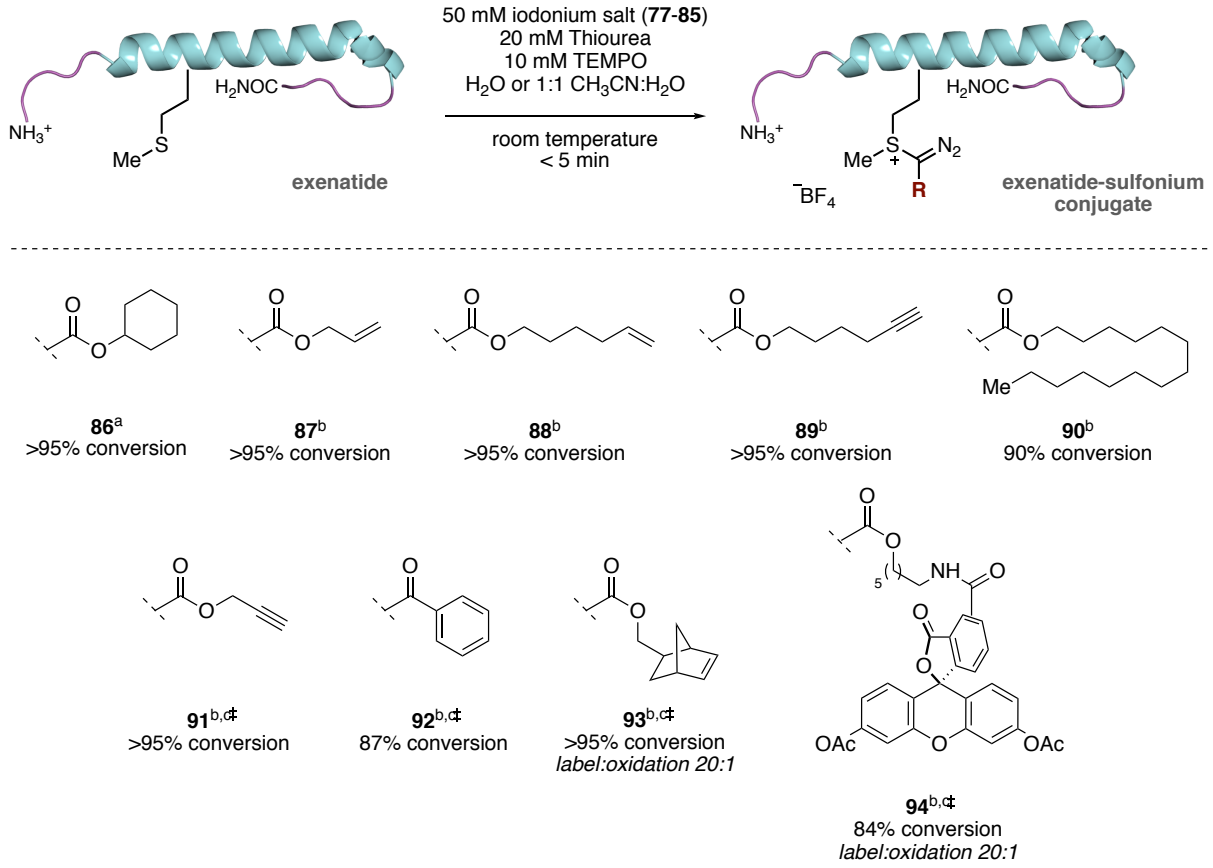
These iodonium salts demonstrated limited stability over prolonged periods at room temperature, however were stable for >3 months when stored as a solid, under air at -20 °C. All iodonium salts required minimal purification, requiring only trituration and washing of the desired product, providing a facile isolation procedure. Iodonium salts **54** and **77** could be isolated in the solid state, yielding a yellow solid which could subsequently be handled with ease. All other iodonium salts resulting from modification of the ester moiety were isolated as a gum but could be readily used without complex purification. The straightforward synthesis and purification of these iodonium salts, alongside their desirable stability, provides a strategy for polypeptide and protein modification which may easily be tailored to incorporate numerous functional groups.



Scheme 47 a) General scheme outlining the synthesis of iodonium salt derivatives altering the ester group, R. b) Scope of iodonium salts. [c] Synthesised directly from α-diazoacetophenone. ‡ Synthesised by Dr M. T. Taylor.

The synthesised iodonium salt reagents were tested in the methionine-selective labelling of exenatide using the previously optimised reaction conditions (Scheme 48). Many functional groups were tolerated well by the reaction, achieving quantitative conversion for the synthesis of sulfonium conjugates **86-89**, **91** and **93**. Conversion to myristyl conjugate **90** proceeded

cleanly, although the reaction did not reach completion, with 10% of starting material exenatide remaining. Phenyl ketone conjugate **92** and fluorescein conjugate **94** could also be formed, albeit with lower conversion.



Scheme 48 Reaction of exenatide with various iodonium salts to form exenatide-sulfonium conjugates bearing a range of functional payloads. Conversion determined *via* LCMS total ion count. ‡ Reaction carried out by Dr M. T. Taylor. [a] H_2O. [b] 1:1 $CH_3CN:H_2O$. [c] 5 mM Formic acid.

2.4.4 Stability studies of methionine sulfonium conjugates

While sulfonium species reported by Deming demonstrated excellent stability,[187] analogous sulfonium species formed through the reaction of CNBr with methionine are known to undergo backbone cleavage (see Section 1.3). It was therefore important to carry out stability studies of novel β-sulfonium α-diazoesters to determine the half-life of these species.

Initially, the stability of the exenatide conjugate bearing an ethyl ester, **56**, was investigated over multiple hours at room temperature, in H_2O (Figure 5). Pleasingly, the sulfonium conjugate exhibited a half-life of approximately 96 hours, similar to that of maleimide-cysteine conjugates within certain pH ranges.[230,231]

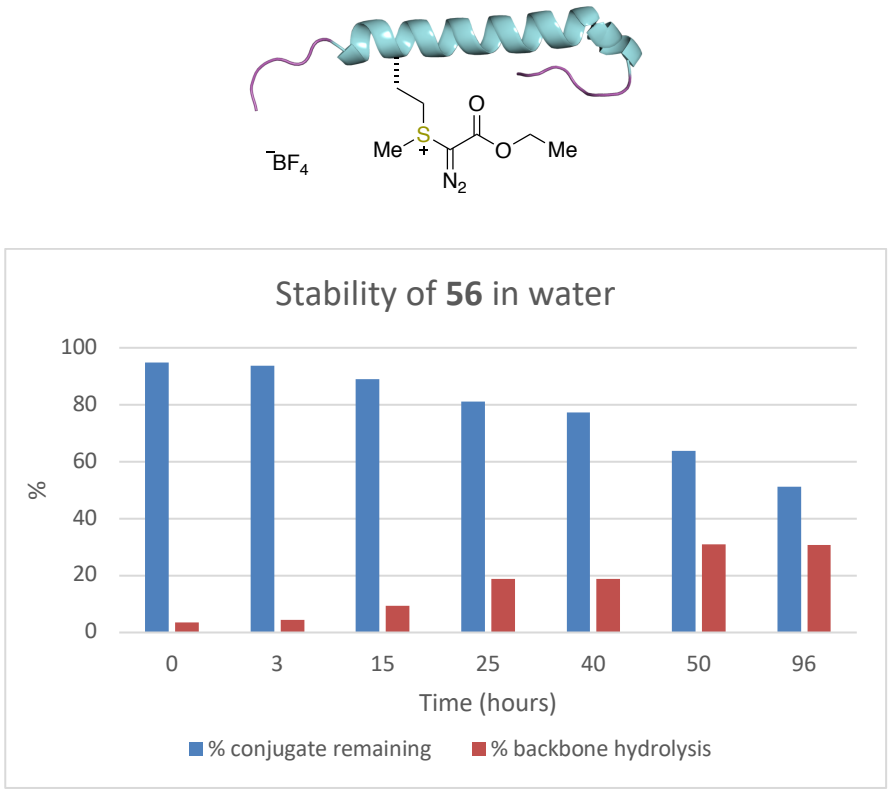
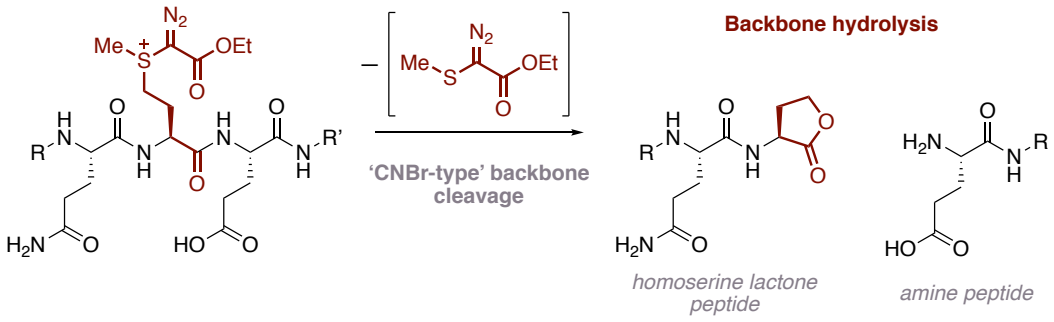


Figure 5 Stability study of exenatide conjugate **56** in H_2O at room temperature. t = 0 hours begins after storage of **56** for 24 hours. Degradation occurs predominantly via backbone hydrolysis.

Degradation of the sulfonium conjugate was predominantly through a 'cyanogen bromide-type' cleavage of the peptide backbone, *via* displacement of neutral thioether by the proximal carbonyl group and subsequent hydrolysis to form homoserine lactone and a corresponding amine peptide fragment (Scheme 49).



Scheme 49 Degradation of exenatide conjugate **56** over time, in H_2O at room temperature, predominantly occurs *via* backbone hydrolysis

An analogous stability study was carried out on the exenatide sulfonium conjugate **86** bearing a cyclohexyl ester (Figure 6). Interestingly, this species showed an improved stability profile, with approximately 70% of the conjugate still remaining after 96 hours, giving a half-life of

approximately 140 hours. This suggests that tailoring the structure of the ester payload may enable the stability of resulting conjugates to be modulated.

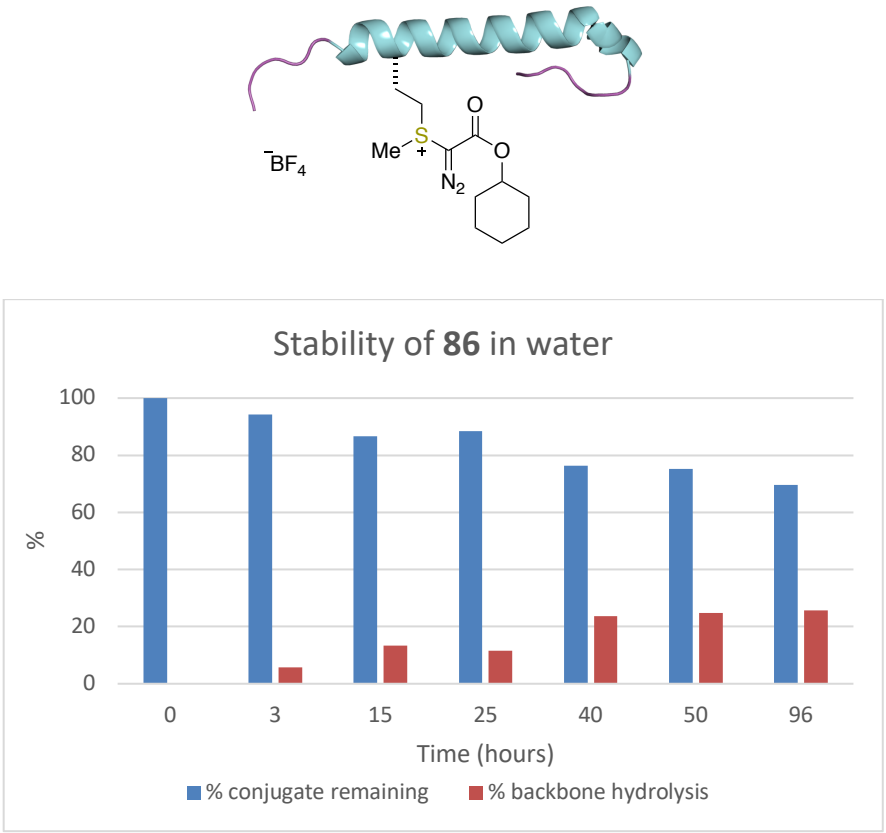


Figure 6 Stability study of exenatide conjugate **86** in H_2O at room temperature

Additional stability studies were carried out by Dr M. T. Taylor in order to examine the behaviour of the diazo sulfonium conjugates in a variety of aqueous media. A solution of propargyl conjugate **91** was monitored at room temperature in H_2O, pH 3 formic acid or pH 7.4 phosphate buffer over a number of hours (Figure 7). In the case of the propargyl ester species, the stability was evidently poorer than the cyclohexyl or ethyl ester analogues, with a half-life of <50 hours at room temperature. The nature of the aqueous media had a dramatic effect on sulfonium stability, with the conjugate proving significantly less stable in basic phosphate buffer compared to water or acidic conditions. In the same way as **56** and **86**, **91** degraded through peptide backbone hydrolysis. However, an additional degradation pathway was observed for **91**; cleavage of the sulfonium label to yield the initial exenatide polypeptide with unmodified methionine residue.

Dr M. T. Taylor also observed partial stability of sulfonium conjugate **91** over 8 hours in the presence of the antioxidant glutathione, commonly found in cells.[33] The primary reaction that

occurred was the reduction of the diazo moiety to the hydrazone. Although technically a degradative pathway, the label remained intact, and no cleavage was observed. While this study does indicate the instability of the sulfonium species to glutathione, it is notable that the payload is not cleaved from the substrate, an important factor for functionalised protein products.

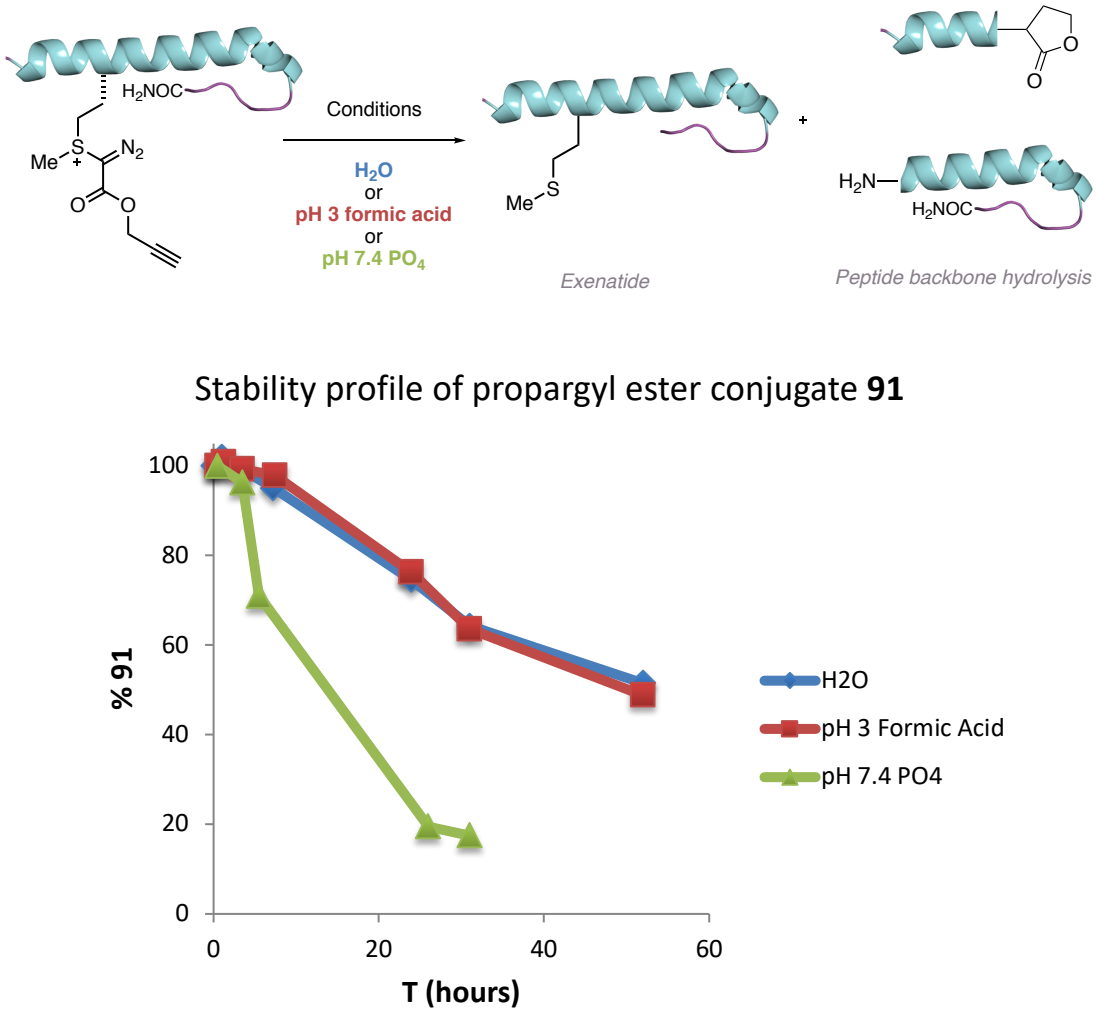
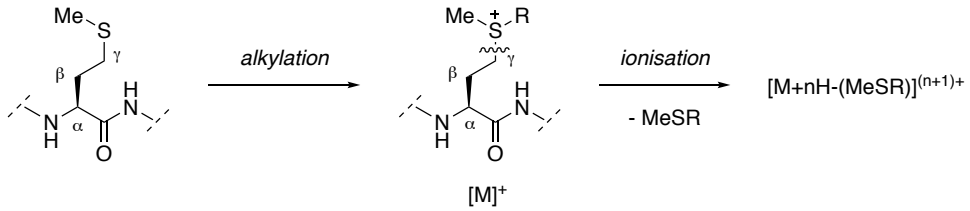


Figure 7 Stability profile of **91**. For each experiment, **91** (100 µM) was allowed to stand in solution at room temperature (21°C) over the time course and analyzed by LCMS. **91** degrades *via* a mixture of peptide backbone hydrolysis and label hydrolysis. Studies carried out by Dr M. T. Taylor.

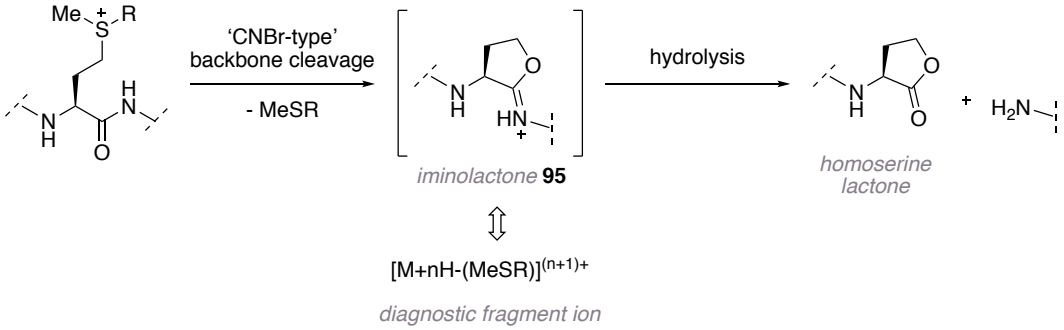
2.4.5 Mass spectral characteristics of β-sulfonium α-diazoester-containing peptide and protein conjugates

A well-established phenomenon within mass spectrometry is that upon ionisation, alkylated methionine residues can fragment at the Cγ-SR_2 bond to produce an $[M\text{-}(SR_2\text{-}H)+H]^+$ ion, in addition to the formation of the parent molecular ion $[M]^+$ (Scheme 50).[188,232,233] This fragmentation pattern is unique to methionine and, therefore, it may be used as a key diagnostic tool to confirm functionalisation at methionine residues.



Scheme 50 Diagnostic fragmentation of alkylated methionine derivatives upon ionisation during mass spectrometry

Iminolactone **95** is an intermediate in the cyanogen bromide-mediated backbone cleavage of alkylated methionine derivatives (Scheme 51).[189,190] This iminolactone is identical in mass to the diagnostic fragment ion observed in the mass spectrum of alkylated methionine derivatives. These iminolactone species are readily hydrolysed under acidic conditions to form the corresponding homoserine lactone peptide and amino peptide. Therefore, it is not expected that this iminolactone species would be observed by mass spectrometry. However, in order to unambiguously assign the fragment ion observed to the diagnostic fragmentation and not the iminolactone degradation product, the origin of this molecular ion was further explored.



Scheme 51 Cyanogen bromide-mediated peptide backbone cleavage. The intermediate iminolactone species has an identical mass to the fragment ion observed upon mass spectral ionisation.

In order to confirm that the β-sulfonium α-diazoester species synthesised displayed the diagnostic fragmentation pattern, isolation and characterisation of dipeptide model system **51** was carried out. The methionine residue was alkylated and the resulting sulfonium species **52**

isolated as a single pure product in >90% purity. **52** clearly displayed the diagnostic fragment ion, $[M\text{-}(MeSR)+H]^+$ (m/z 159.0760) in addition to the molecular ion, $[M]^+$ (m/z 319.1061). This demonstrated that the fragment was created upon ionisation in the mass spectrometer and was not the result of CNBr-type degradation of the isolated sample (Figure 8).

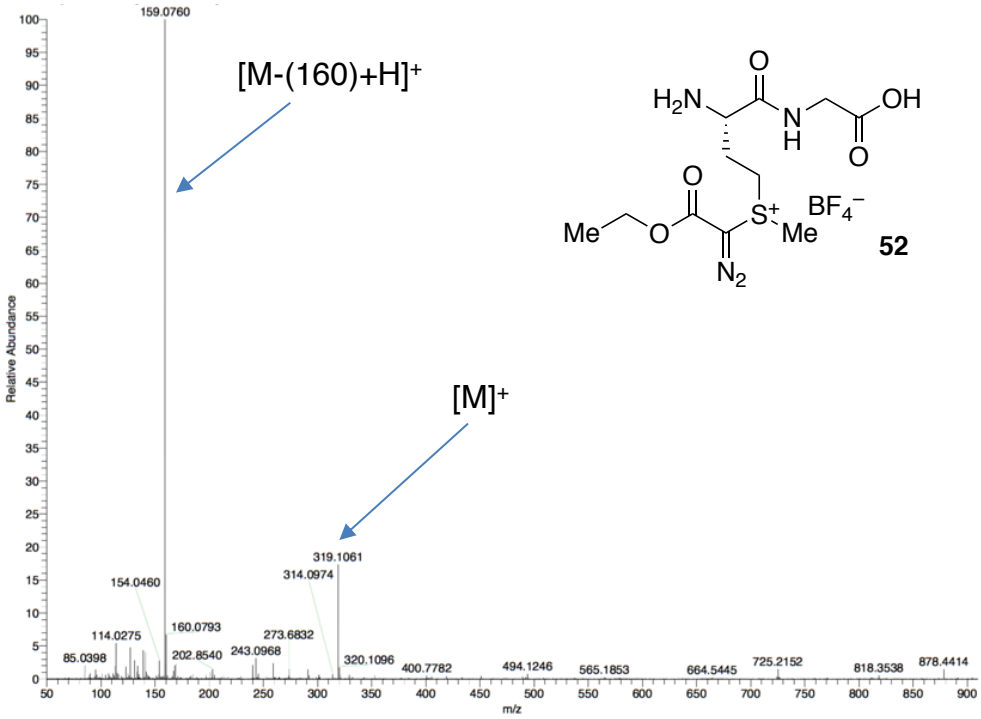


Figure 8 Mass spectrum of MetGly sulfonium conjugate **52**, clearly showing both the molecular ion $[M]^+$ and the diagnostic fragment ion $[M\text{-}(160)+H]^+$.

Next, the mass traces of two different exenatide sulfonium species, **56** and **90**, were considered, bearing an ethyl diazoester and a myristyl diazoester respectively (Figure 9). Due to the non-polar, aliphatic chain, **90** eluted at a much later retention time (~3.2 min) compared to **56** (~2.2 min). However, an identical fragment mass was observed (e.g. M = 1035.5) in both mass chromatograms. Were these ions a result of cyanogen bromide-type degradation, the iminolactone side-product would be identical and would appear at the same retention time for both species. As this was not the case, it was concluded that the ion resulted from the expected diagnostic fragmentation.

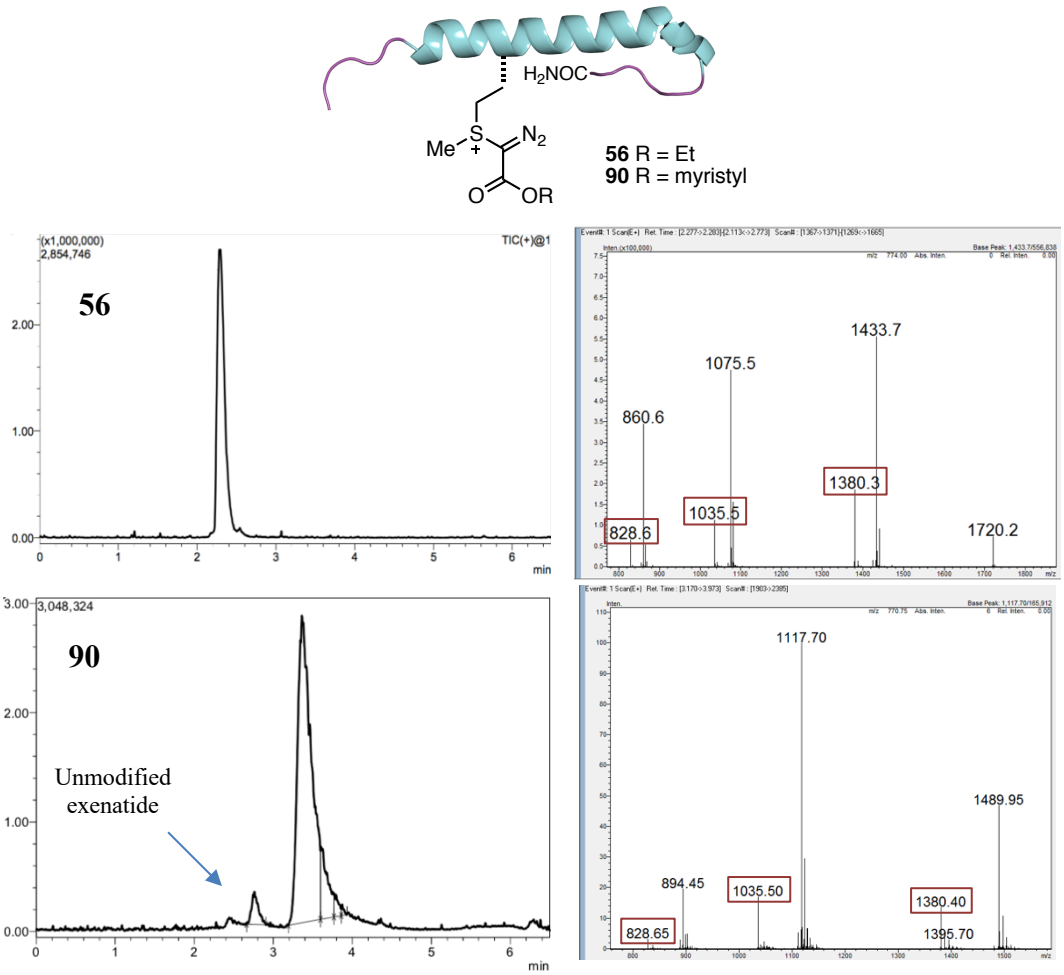
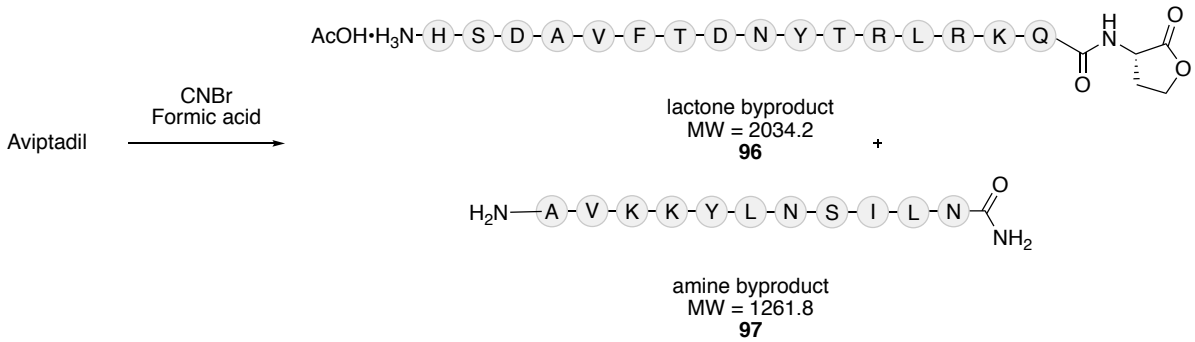


Figure 9 Comparison between the mass spectra of exenatide conjugates **56** and **90** bearing differing ester groups. As a result of the ester groups, the two sulfonium conjugates elute from the LCMS at different retention times. However, both display identical fragmentation patterns.

In an attempt to observe an authentic spectrum of the putative iminolactone intermediate **95**, the polypeptide aviptadil was stirred with cyanogen bromide under acidic conditions. The iminolactone was not observed, with only the corresponding hydrolysed peptidic fragments **96** and **97** detected clearly by LCMS at different retention times (homoserine lactone **96** and amine fragment **97**, Scheme 52). The absence of the iminolactone indicated the instability of this intermediate to hydrolysis and hence provided further evidence that the iminolactone was not responsible for the observed ion within the mass spectrum of the sulfonium conjugates.



Scheme 52 Cyanogen bromide-mediated cleavage of polypeptide aviptadil, forming homoserine lactone peptide **96** and amine peptide **97**. No iminolactone intermediate could be observed.

2.4.6 Exploiting the reactivity of the diazo motif in β-sulfonium α-diazoester conjugates

The novel alkylative bioconjugation reaction developed herein introduces a non-native, high energy diazo motif. This newly introduced, protein-bound bioorthogonal handle presents an opportunity for further exploitation of resulting conjugates in subsequent transformations. Diazo motifs are known to display a diverse reactivity manifold, with different reagents and catalysts able to induce distinct chemical transformations (Figure 10).

Nucleophilic attack of phosphine reagents at the terminal nitrogen atom of diazo species is precedented, inducing Staudinger-like formation of a phosphine-imine bond (Figure 10a).[234] Additionally, thermal or photochemical generation of metal carbenoid species using transition metals has been extensively studied (Figure 10b).[235] Photochemical reactions of diazo species are also well known, through photoinitiated formation of carbenes[236] or carbon-centred radicals (Figure 10c)[237,238] which are subsequently employed in various transformations. As such, it was envisaged that the β-sulfonium α-diazoester would display analogous reactivity to that of related diazo species, allowing further diversification of protein conjugates.

Investigation into Staudinger-inspired reactivity of the β-sulfonium α-diazoester conjugates and the visible light-induced formation of carbon-centred radicals is described herein. However, efforts to explore the formation and reactivity of metal carbenoids derived from the β-sulfonium α-diazoester conjugate are ongoing within the group and are beyond the scope of this report.

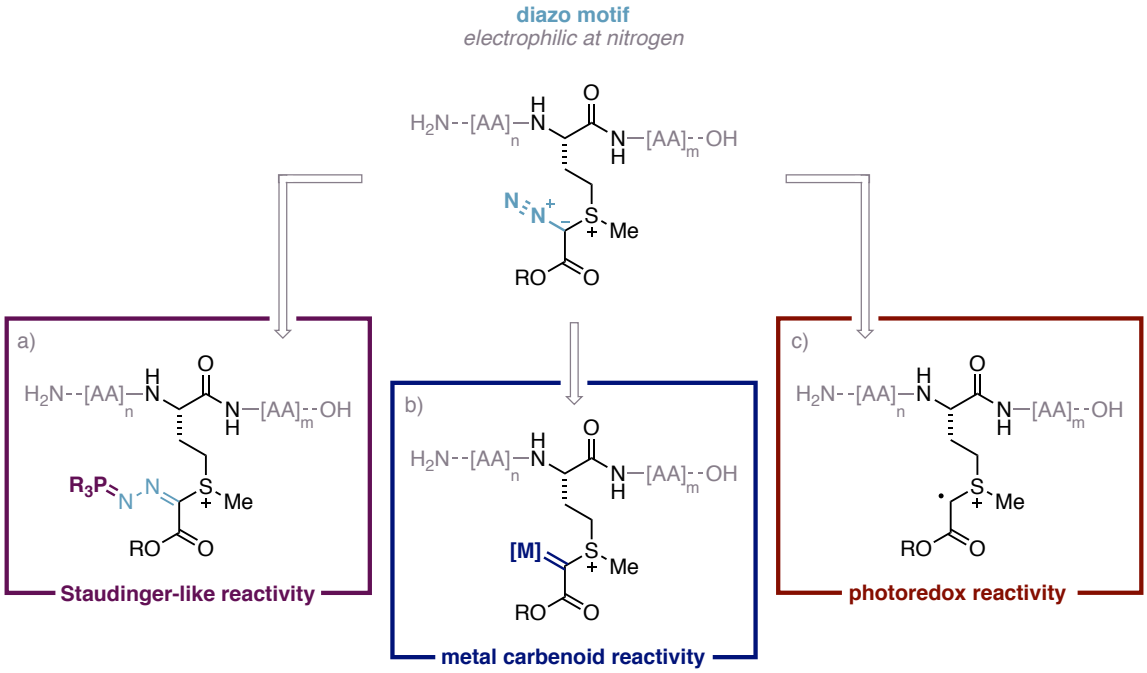


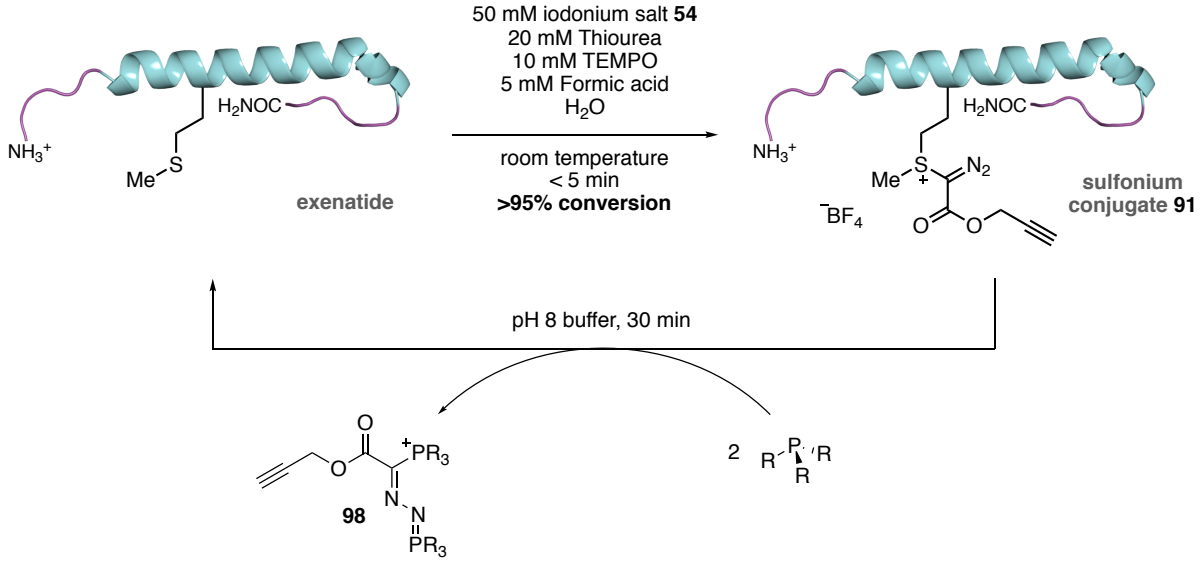
Figure 10 Distinct modes of reactivity of diazo functional groups through the use of varied reagents and catalysts

2.4.6.1 Phosphine-promoted cleavage of methionine label

Reversible modification of polypeptides through triggered release of payloads is advantageous for many biological applications, for example the release of therapeutic peptides from a carrier or recovery of affinity purified peptides for proteomic analysis.[239,240] As a potential tool for the triggered release of polypeptides, the reversibility of the methionine bioconjugation was investigated. As previously discussed, Deming reported the general, triggered reversal of their methionine alkylation strategy using 2-mercaptopyridine as a nucleophilic cleavage reagent[187] (Section 1.3). Inspired by this, we turned our attention to investigating whether successful cleavage of the diazoester moiety could be achieved.

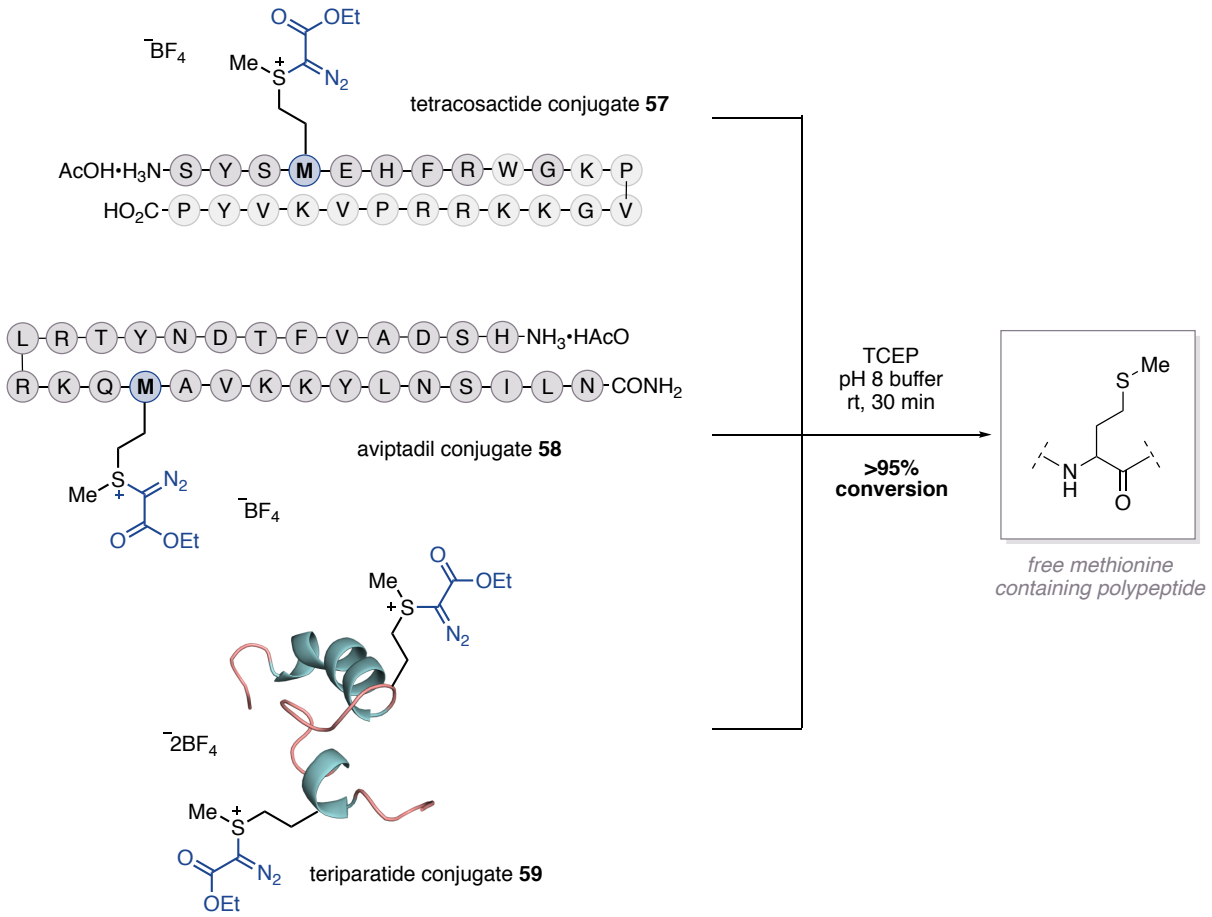
Dr M. T. Taylor screened a number of phosphine reagents under varying reaction conditions in the hope of achieving reversal of the labelling with Staudinger-inspired reactivity. It was discovered that nucleophilic phosphine reagents could attack the electrophilic nitrogen atom of the diazo motif in a similar fashion to the Staudinger reaction, forming a phosphine-imine bond. This could be followed by a second molecule of the phosphine reagent displacing the sulfonium species, regenerating the neutral thioether of exenatide and releasing cationic phosphorus species **98** (Scheme 53). The most successful reagents for the triggered cleavage of the

methionine label were found to be triphenyl phosphine or tris(2-carboxyethyl)phosphine (TCEP), achieving quantitative conversion over 30 minutes.



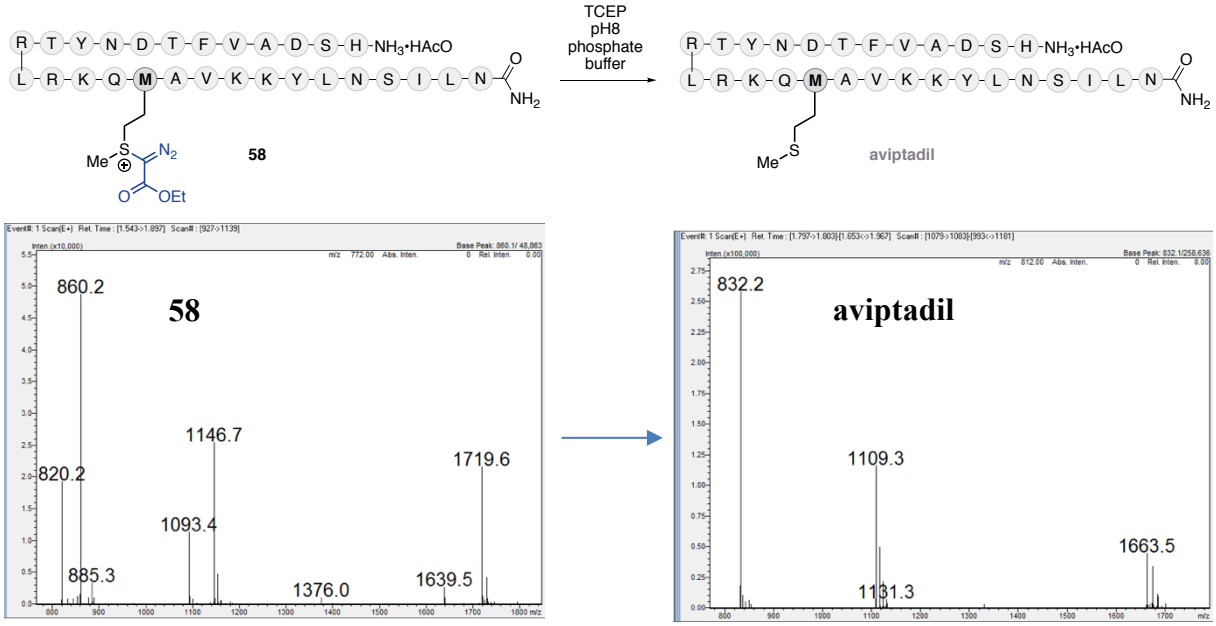
Scheme 53 Triggered cleavage of methionine diazoester label using phosphine reagents (e.g. PPh_3 or TCEP), reforming the free polypeptide exenatide. Conditions developed by Dr M. T. Taylor.

With optimised conditions in hand, the reversal of the labelling was investigated on the scope of polypeptide substrates. In all cases, >95% conversion to unmodified polypeptide was observed in 30 minutes (Scheme 54). For the teriparatide bis-sulfonium conjugate **59**, cleavage of the label was achieved at both methionine residues, with no evidence of the singly-cleaved product. In order for the release of the label to occur cleanly, purified sulfonium polypeptide product must be used in the reaction. When the reaction was attempted on the crude sulfonium species, unmodified product was observed in much lower conversion and the reaction profile proved complex. However, Dr M. T. Taylor demonstrated that the label of unpurified protein sulfonium conjugate ubiquitin **63** could be cleaved. Upon treatment with TCEP, **63** was reverted back to free ubiquitin in >90% conversion.



Scheme 54 Triggered reversal of polypeptide conjugates of tetracosactide, aviptadil and teriparatide, regenerating the parent polypeptide with free methionine residue

Interestingly, on consideration of the mass spectrum of aviptadil conjugate **58**, it was observed that, upon hydrolysis, the characteristic $[M-(MeSR)+H]^+$ fragment ion disappeared completely from the mass spectrum (Scheme 55). This observation provided further evidence to corroborate the diagnostic fragmentation of the Cγ-SR_2 bond, discussed in Section 2.4.5. If the ion were a result of peptide backbone cleavage, the iminolactone would be irreversibly hydrolysed to form two smaller peptidic fragments which would remain unaffected by the reduction and would be visible in the mass spectrum. Instead, complete disappearance of the ion was observed, with reformation of the parent polypeptide, again confirming the nature of this diagnostic fragment ion.



Scheme 55 Mass traces of aviptadil conjugate **58**, clearly showing the diagnostic fragment ions (M = 820.2, 1093.4, 1639.5), and the crude reaction of **58** with TCEP, cleaving the sulfonium label and regenerating free aviptadil. No evidence of the fragment ions could be observed after the reversal of the labelling.

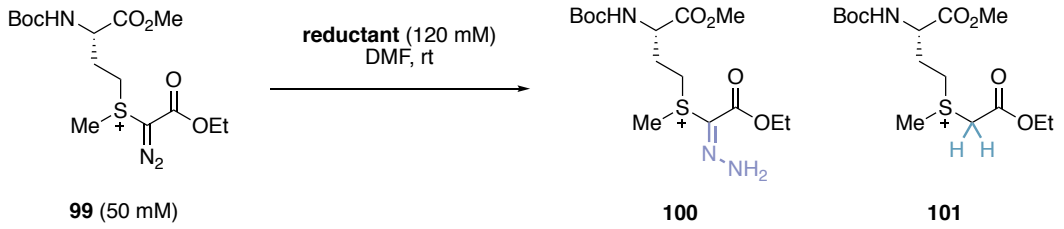
2.4.6.2 Photochemical reduction of α-sulfonium diazoesters

Improved stability of the sulfonium conjugates was desired as it would enable more facile handling of resulting functionalised species and facilitate the use of these conjugates in subsequent bioorthogonal transformations and its use in biological applications. Therefore, efforts focused on improving the stability of β-sulfonium α-diazoester conjugates.

The reduction of the diazo moiety in the sulfonium products, to a methylene unit, would form the corresponding trialkylsulfonium species. It was hypothesised that trialkyl sulfonium derivatives of methionine would likely exhibit improved stability, analogous to those reported in previous methionine alkylation strategies (see Section 1.3).

A screen of reductive conditions was carried out by Dr M. T. Taylor (Scheme 56). Hantzsch ester, **102**,[241] triaryl phosphine **103**[242] and glutathione **104**[243] were screened on a small molecule diazosulfonium species **99** in DMF, at room temperature. While in the case of Hantzsch ester **102** an appreciable amount of the reduced product **101** was observed, under all the other attempted reaction conditions the predominant product was the partially reduced hydrazone **100**.

a)

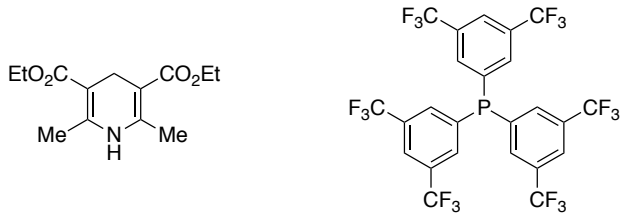


99 (50 mM) **100** **101**

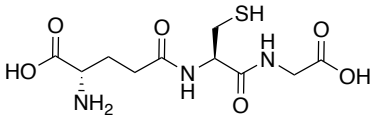
b)

Entry	reductant	ratio **100**:**101**
1	**102**	1:1.5
2	**103**	1:0
3	**104**	1:0

c) **Reductant =**



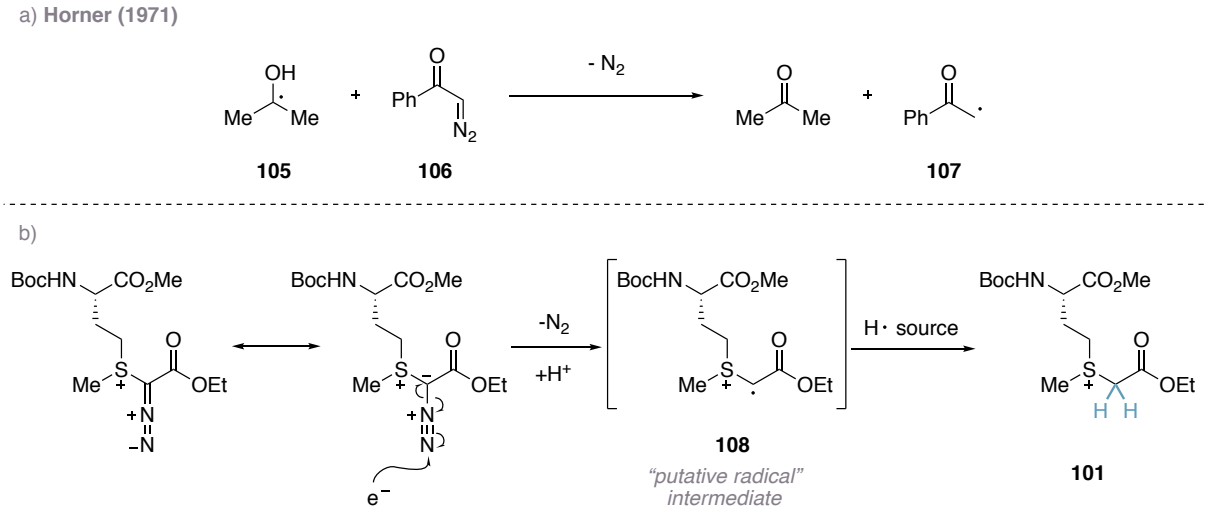
102 **103**



104

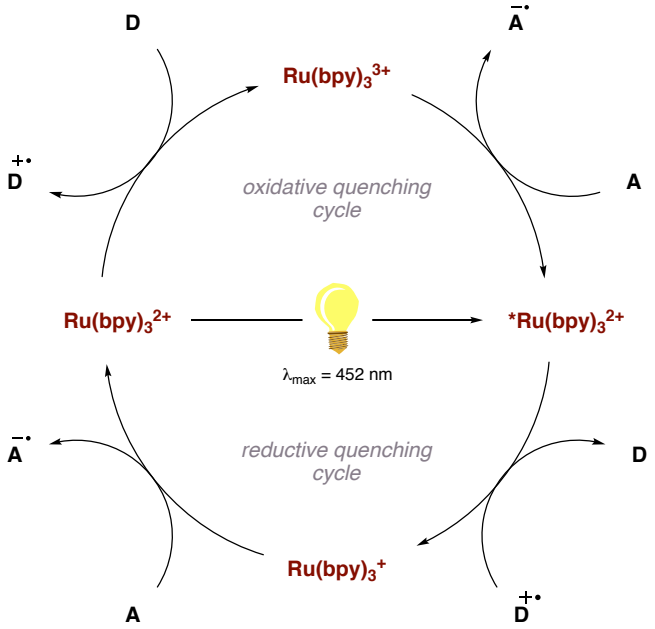
Scheme 56 a) Scheme for the screening of reduction conditions for diazosulfonium species **99**. b) Table for the screening of different reducing agents. c) Reducing agents used in the screening. Reactions carried out by Dr M. T. Taylor.

It was proposed that instead of a 2-electron reduction procedure, a single-electron reduction could provide a mild and selective method for formation of **101**. Radical reactions of α-diazo compounds have been reported, resulting in the elimination of nitrogen and the formation of carbon-centred radicals.[237,238] Horner and co-workers demonstrated that hydroxyalkyl radicals **105** can react with α-diazoketones **106**, causing overall hydrogen atom transfer and loss of dinitrogen, forming carbon-centred radicals **107** (Scheme 57a).[238] They proposed that initial homolytic addition took place at the terminal nitrogen atom of the α-diazo moiety. It was therefore postulated that an analogous single electron reduction could be developed at the terminal nitrogen of the diazo moiety. Through a similar hydrogen atom transfer and nitrogen extrusion pathway, a putative carbon-centred radical may be formed **108** (Scheme 57b). Subsequent transfer of a hydrogen atom may yield the desired reduction product **101**.



Scheme 57 a) Reaction of α-diazoketones with hydroxylalkyl radicals to form carbon-centred radicals. b) Schematic of the proposed reactivity of diazosulfonium species under single-electron reduction conditions

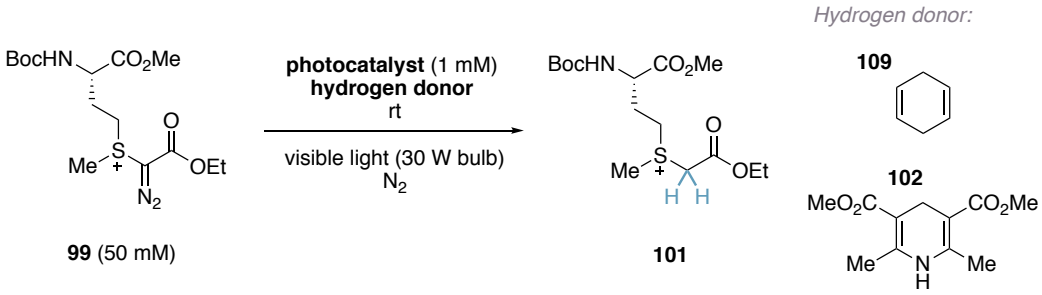
It was envisaged that this single electron reduction could be achieved using a photocatalytic approach. Photocatalysis relies on the use of visible light to excite a photocatalyst, typically a metal complex or organic species. In its excited state, the photocatalyst has an increased propensity to act as both an electron donor and acceptor, through either a reductive or oxidative quenching cycle (Scheme 58). Therefore, visible light photocatalysis offers access to both a powerfully oxidising and reducing species under mild conditions.



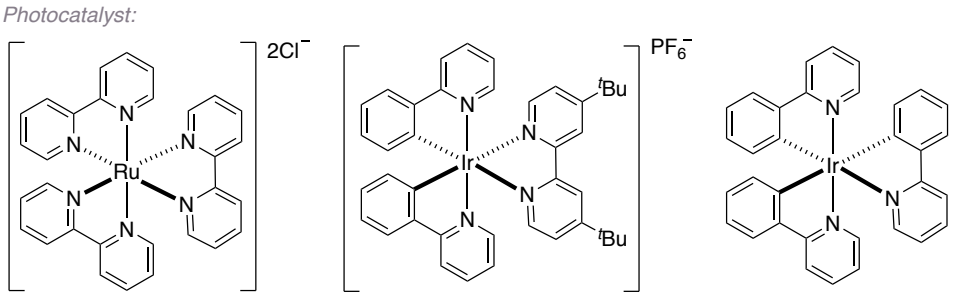
Scheme 58 Oxidative and reductive quenching cycle of $Ru(bpy)_3$. Upon excitation with visible light, the excited photocatalyst can proceed by either oxidative or reductive quenching. In the oxidative quenching cycle, the excited photocatalyst acts as a single electron reductant of an electron acceptor, A. In the reductive quenching cycle, the excited photocatalyst acts as a single electron oxidant of an electron donor, D.[244]

Screening of various photocatalytic conditions was carried out by Dr M. T. Taylor under visible light irradiation (Table 3). A hydrogen atom transfer donor (HAT donor) would be required in order to trap the proposed radical forming as a result of single-electron reduction. 1,4-cyclohexadiene **109** and Hantzsch ester **102** are both known HAT sources[245,246] and were the two HAT sources used for all of the screening reactions. A range of photocatalysts were screened by Dr M. T. Taylor, with reduction potentials in the range -0.81 to -1.73 eV.[244] Conversion to the reduced sulfonium product **101** correlated with the reduction potentials of the photocatalysts screened (Table 3, Entries 1-4), with the most strongly reducing photocatalyst, $Ir(ppy)_3$, achieving the highest conversion to product (Table 3, Entry 4). Additional screening was carried out by Dr M. T. Taylor and it was found that >95% conversion to reduction product **101** could be achieved using Hantzsch ester **102** as the HAT source (Entry 5, Table 3). Pleasingly, using 1:1 CH_3CN:H_2O as the solvent provided a more biocompatible reaction media, without decreasing the conversion to the desired product.

Table 3 Optimisation of reduction procedure for diazosulfonium species using various photocatalysts. All reactions carried out by Dr M. T. Taylor.



Entry	photocat.	H donor	[H donor]	Solvent	%conv	**%101**	Comment
1	—	**109**	10 mM	CH_3CN	32	**1**	—
2	$Ru(bpy)_3Cl_2$	**109**	10 mM	CH_3CN	48	**10**	—
3	$Ir[(dtbbpy)(ppy)_2]^+$	**109**	10 mM	CH_3CN	86	**22**	—
4	$Ir(ppy)_3$	**109**	10 mM	CH_3CN	83	**35**	—
5	$Ir(ppy)_3$	**102**	15 mM	1:1 CH_3CN:H_2O	>95	**95**	500 μM photocat.

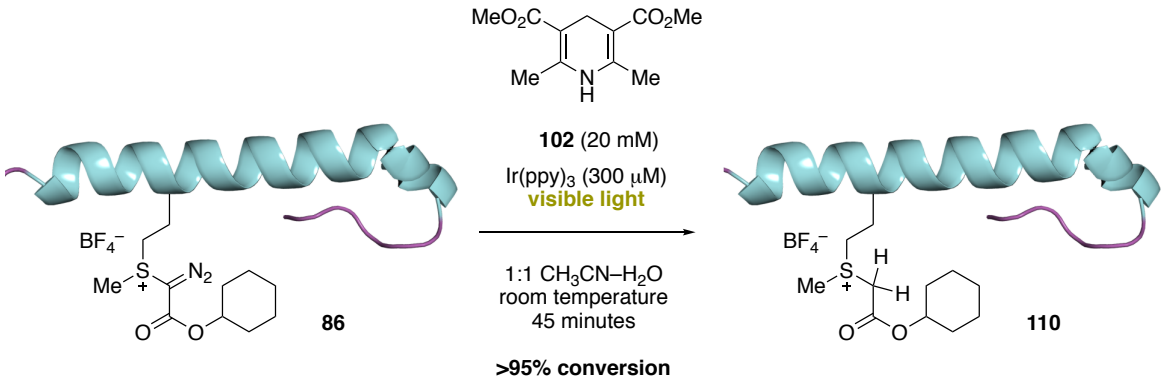


$Ru(bpy)_3Cl_2$
$E_{1/2}$ = -0.81 eV

$[Ir(dtbbpy)(ppy)_2]PF_6$
$E_{1/2}$ = -0.96 eV

$Ir(ppy)_3$
$E_{1/2}$ = -1.73 eV

With optimised conditions in hand for the model sulfonium system, the aim was to translate the photochemical reduction to polypeptide and protein systems. It was found that the optimised conditions were applicable to exenatide sulfonium conjugate **86**, with >95% conversion achieved in 45 minutes (Scheme 59). The reduction was carried out on crude sulfonium conjugate **86** and did not require purification of the diazosulfonium species prior to reduction. As a result, the labelling and photoreduction could be carried out in a two-step one-pot procedure to rapidly furnish the desired trialkylsulfonium derivative **110** from the native polypeptide. Remarkably, no other functionality present in exenatide was reduced using this method, highlighting the mild and selective nature of the photoredox-mediated reduction.



Scheme 59 Optimised conditions for the photochemical reduction of sulfonium diazoesters

Following the reduction of the diazo sulfonium species, investigation into the stability of the reduced trialkylsulfonium species was carried out. The aim was to observe the effect of the reduction on the stability of the resulting products in comparison to the diazosulfonium analogues. Exenatide trialkylsulfonium conjugate **110** proved indefinitely stable towards various concentrations of glutathione (Figure 11), a significant improvement compared to the diazosulfonium species.

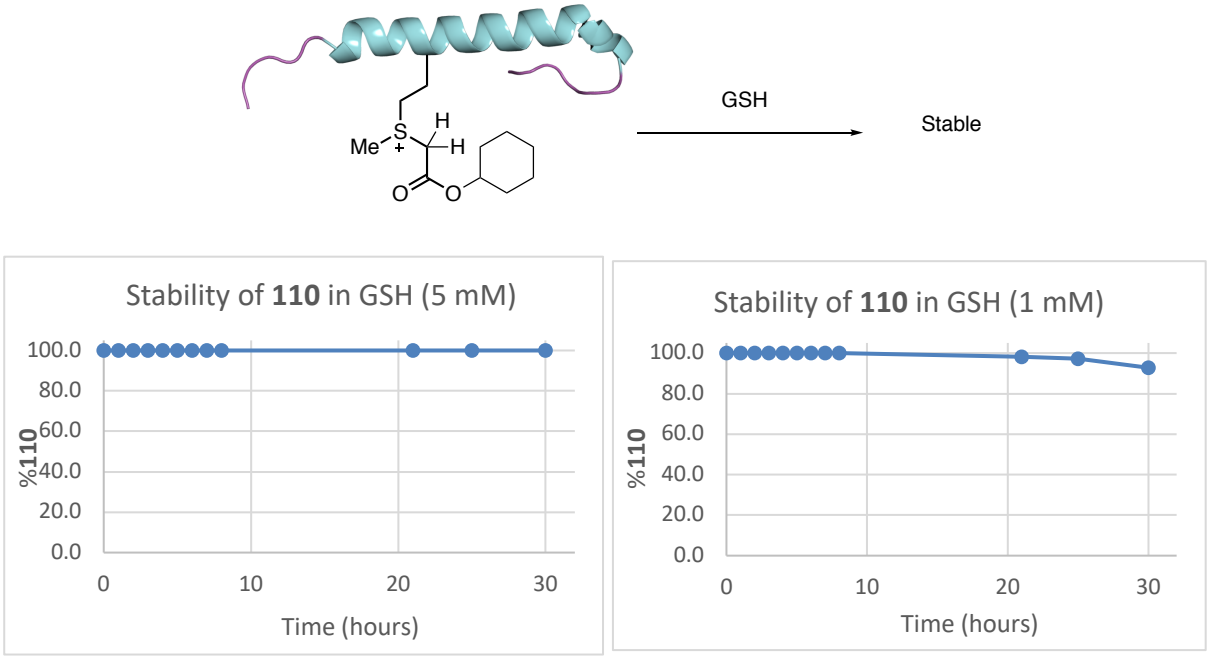


Figure 11 Stability study of exenatide trialkylsulfonium species **110** towards glutathione. **110** showed no signs of degradation over multiple hours.

Additional stability studies were carried out by Dr M. T. Taylor for glucagon trialkylsulfonium conjugate **111** in a variety of aqueous media (Figure 12). In all cases, degradation occurred exclusively through ester hydrolysis to form the carboxylate **112**. While stability in pH 7.4

phosphate buffer was still poor, greatly improved stability was observed in water and pH 3 buffer.

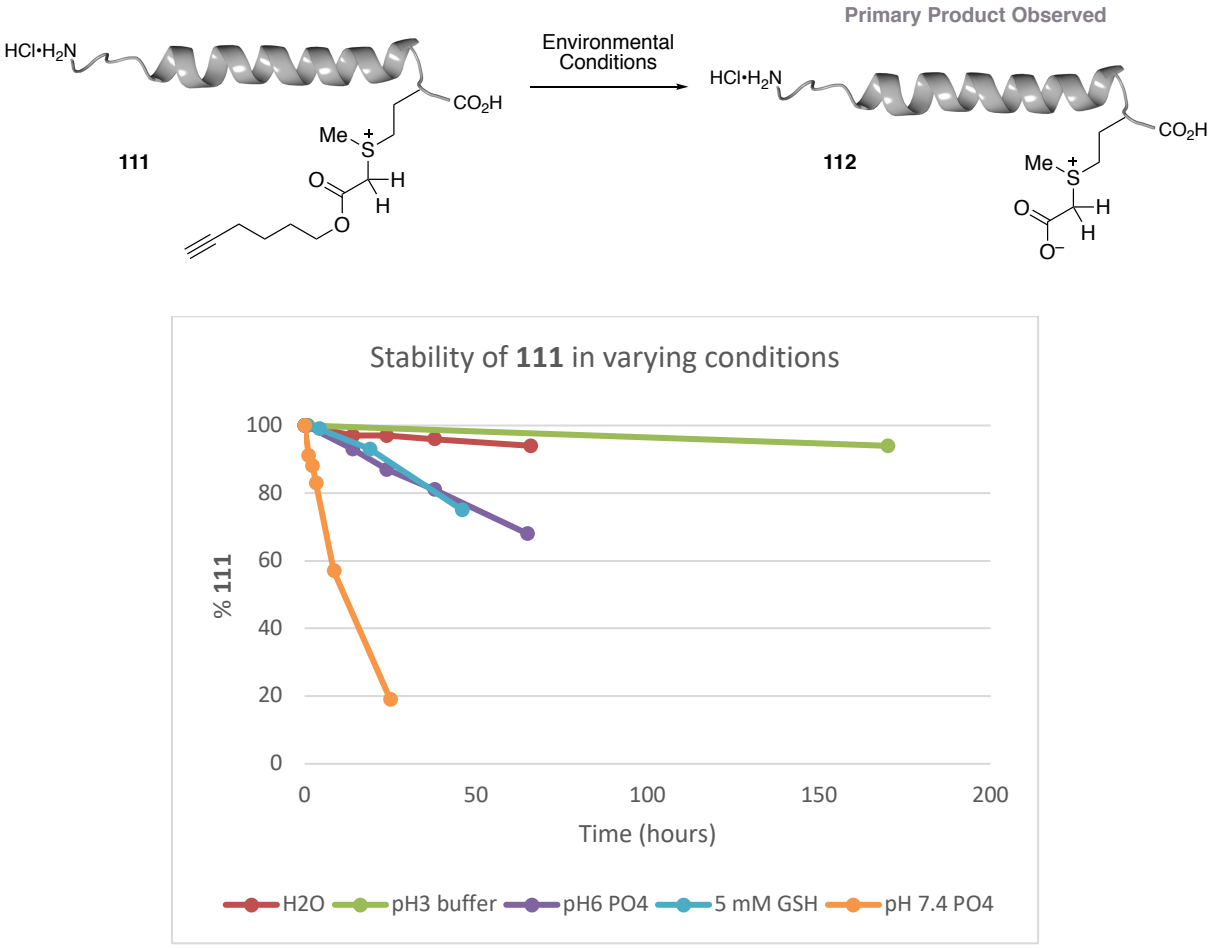
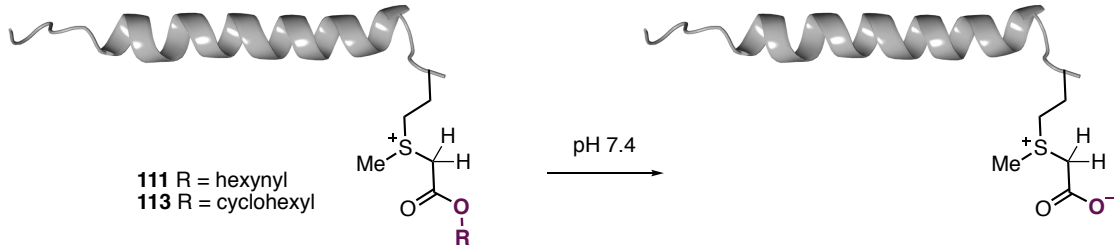


Figure 12 Stability profile of **111**. In all instances, degradation of **111** occurs via the hydrolysis of the ester moiety to yield a free carboxylate. Studies carried out by Dr M. T. Taylor.

Dr M. T. Taylor was also able to observe the stabilities of trialkylsulfonium conjugates of the polypeptide glucagon with two different ester moieties **111** and **113** (Figure 13). The cyclohexyl ester exhibited a significantly higher stability than the corresponding hexynyl ester, with a half-life of approximately 90 hours compared to only 10 hours for the hexynyl conjugate. In both cases decomposition occurred *via* ester hydrolysis, to yield the carboxylate. It was hypothesised that the increased steric hindrance of the cyclohexyl ester could disfavour ester hydrolysis, resulting in superior stability.



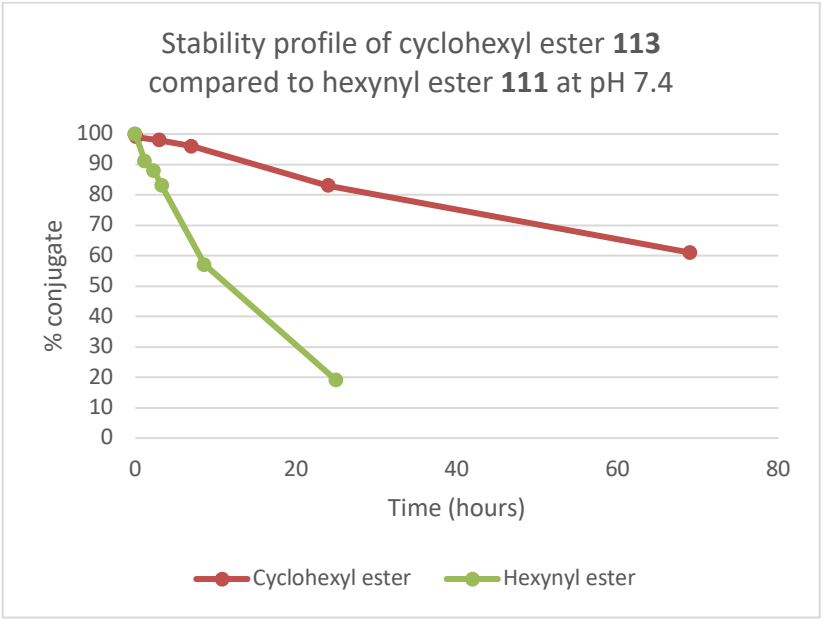
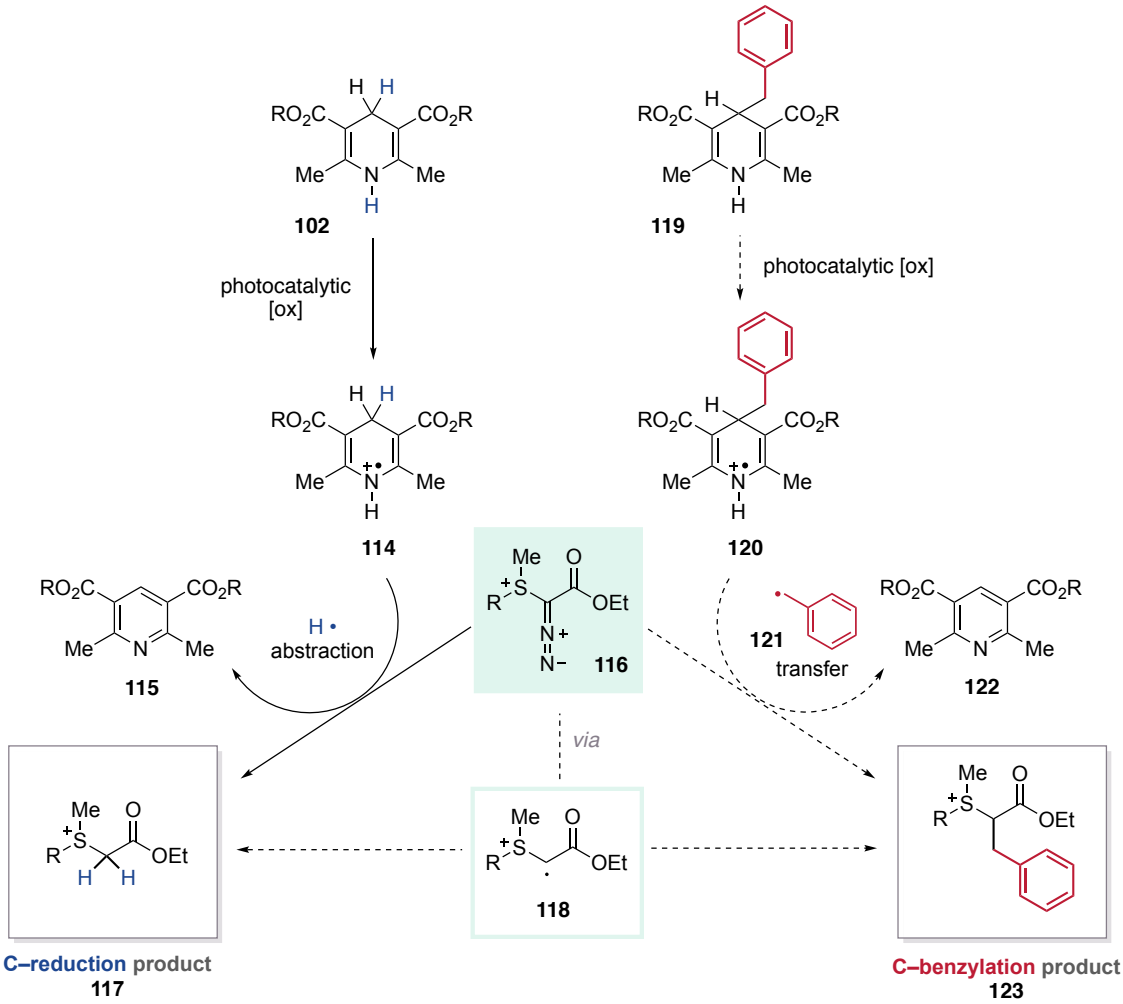


Figure 13 Comparative stability studies of glucagon trialkylsulfonium conjugates **111** and **113** observing the effect of the ester group on the stability of the species in pH 7.4 buffer. Reactions carried out by Dr M. T. Taylor.

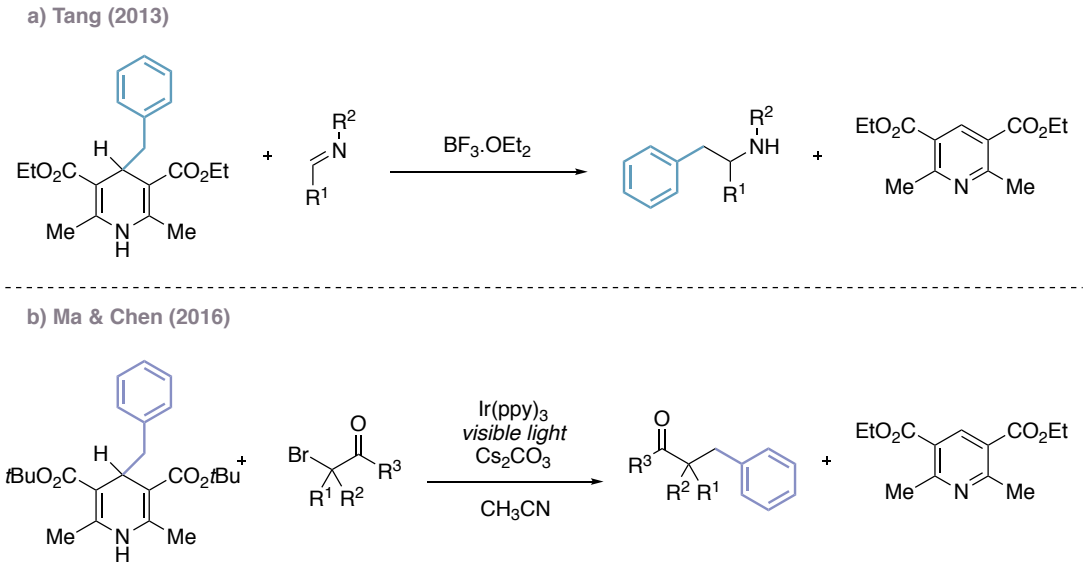
2.4.6.3 A photoredox-mediated bioorthogonal transformation at β-sulfonium α-diazoesters

Following the development of a mild and selective photochemical reduction of the β-sulfonium α-diazoester moiety, it was hoped that this highly reactive diazo motif could be exploited as a platform for subsequent bioorthogonal transformations to introduce additional functionality of biological relevance. It had been demonstrated previously that Hantzsch esters could be used as hydrogen atom donors[247] in the photoredox-mediated reduction of sulfonium diazoesters, proposed to be occurring *via* a putative radical intermediate **118** (Section 2.4.6.2). It was hypothesised that, by analogy, the use of a C4-substituted Hantzsch ester **119** (Scheme 60) and a photocatalyst would enable the generation of a benzyl radical **121** which may be formally transferred to the α-position of the diazoester **116** through coupling with putative carbon-centred radical **118**. If successful, modification of the diazo moiety could be achieved, introducing additional exogenous functionality using a photochemical-mediated bioorthogonal transformation.



Scheme 60 Schematic design plan for secondary functionalisation of the diazosulfonium conjugate. Left; photocatalytic oxidation of Hantzsch ester to yield radical cation as a HAT source forming reduction product **117**. Right; C-4 substituted Hantzsch esters as a source of a benzyl radical, a hypothesised radical-functionalisation to form **123**.

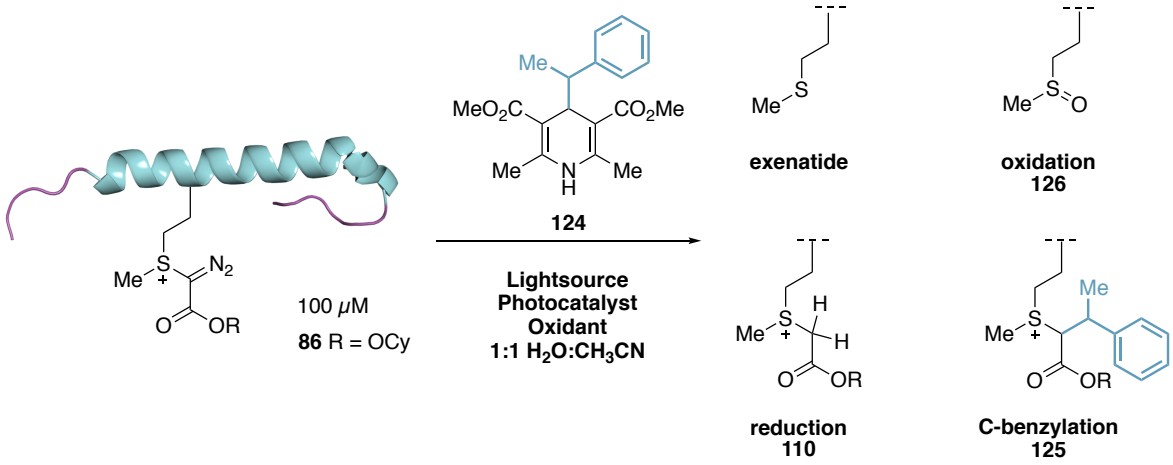
Similar transformations have been reported for the use of C4-substituted Hantzsch esters as alkyl donors[248] using either a Lewis acid,[249] or photoredox catalysis.[250] Hantzsch esters substituted with a benzyl group have been shown to generate a benzyl radical, which can react through radical addition into imines or through coupling with other carbon-centred radicals to form benzyl-substituted products (Scheme 61a, b).[248] Generation of the alkyl radical is reported to require a photocatalyst due to the substituted Hantzsch ester likely absorbing at a wavelength outside the visible region.[248]



Scheme 61 Literature precedence for substituted Hantzsch ester reagents acting as alkyl donors under Lewis acidic conditions or photoredox catalysis.

Initial efforts for the development of this photoredox benzylation process were carried out by Dr M. T. Taylor on model polypeptide sulfonium conjugate exenatide **86**, using 1-phenylethyl substituted Hanztsch ester **124**. The highest conversion to product achieved from the initial screening was 38% (Entry 1, Table 4), through the use of a $Ru(bpy)_3Cl_2$ photocatalyst and a strong oxidant potassium persulfate. The main deleterious pathway was discovered to be the formation of the reduction product **110**, assumed to be formed through abstraction of a hydrogen atom instead of the proposed radical-radical coupling. Cleavage of the methionine label, regenerating free exenatide was also observed, alongside an oxidative side-product **126**, resulting from subsequent oxidation of free methionine to the sulfoxide. It was noted that a more powerful 40 W Kessil light bulb proved more efficient in the reaction compared to a 30 W bulb, allowing a shorter reaction time. This initial hit provided a promising preliminary result that the photoredox benzylation could be carried out on a complex polypeptide system.

Table 4 Preliminary optimisation of photochemical benzylation. All reactions carried out by Dr M. T. Taylor.

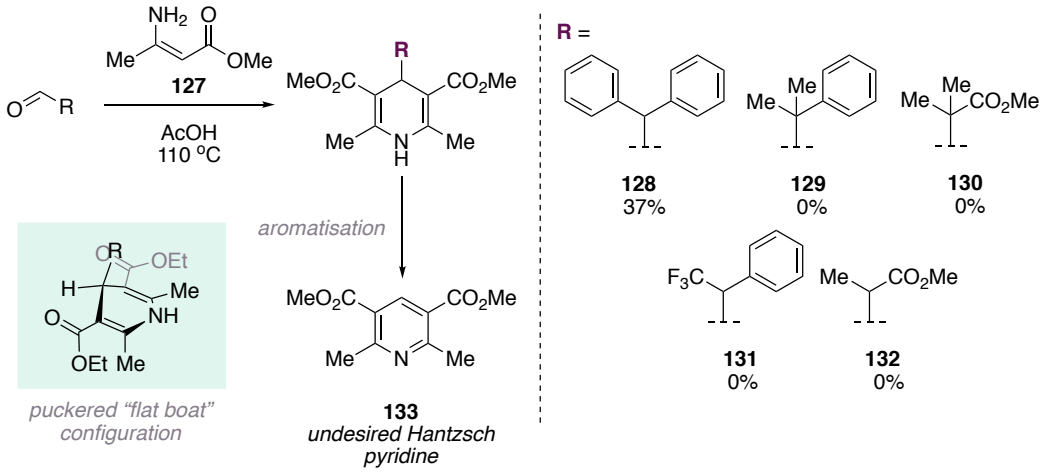


Entry	[Hantzsch]	Photocatalyst	Oxidant	Light[a]	T(min)	%conversion of 86	% C-benzyl[n] 125	% Free peptide	%oxidized peptide	% 110
1[b]	40 mM	$Ru(bpy)_3$ (200 μM)	$K_2S_2O_8$ (10 mM)	40W Kessil	5	94	38	11	7	28

[a]Light placed at 5 cm distance from reaction vessel. [b]Reaction outcome determined by TIC.

Following the identification of an initial hit using a $Ru(bpy)_3Cl_2$ photocatalyst and an oxidant $K_2S_2O_8$, capable of single electron oxidation, it was proposed that modification of the Hantzsch ester reagents may allow modulation of the benzylation reaction. Therefore, the synthesis of alternative Hantzsch esters which would form more stabilised secondary or tertiary benzyl radical species was investigated. Incorporation of a diphenyl substituent proved successful, with Hantzsch ester **128** isolated in 37% yield from diphenylacetaldehyde and methyl-3-aminocrotonate **127** (Scheme 62). Unfortunately, synthesis of a Hantzsch ester bearing a dimethylphenyl substituent **129** was unsuccessful. Using 2-methyl-2-phenylpropanal and methyl-3-aminocrotonate **127**, the corresponding Hantzsch ester **129** could be formed under the reaction conditions but was unstable to purification. Facile dealkylation was observed to occur, forming the aromatised Hantzsch pyridine **133**. After extensive efforts, only the Hantzsch pyridine **133** could be isolated as a result of oxidative aromatisation and dealkylation. The same fate was observed for Hantzsch ester **130**, with a quaternary centre. Similarly, Hantzsch esters **131** and **132**, bearing tertiary carbon substituents, could not be synthesised, presumably due to similar instability of the products. Correlations could be drawn between the stability of the radical formed and the tendency for aromatisation to the Hantzsch pyridine **133**. Additionally, C4-substituted Hantzsch esters are known to adopt a higher energy, 'puckered' flat boat conformation (Scheme 62).[251] As such, it was thought that steric factors of the more highly

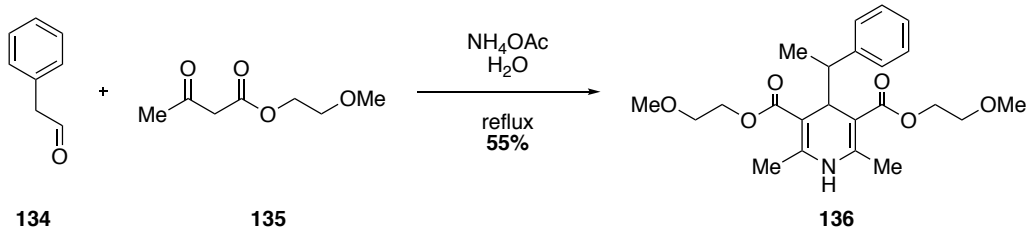
substituted Hantzsch esters, bearing quaternary substituted carbon centres, may be negatively affecting the stability through further destabilisation of the flat boat conformation.



Scheme 62 Scheme for the synthesis of Hantzsch ester derivatives **128-132**

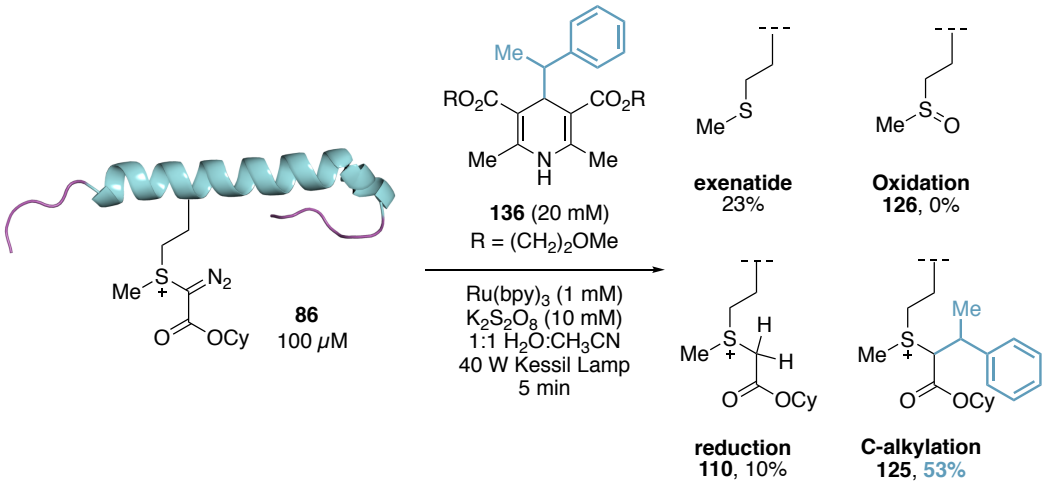
Benzhydryl Hantzsch ester **128** was screened extensively under various photoredox benzylation conditions, although no evidence of benzylation product was observed by LCMS analysis. The predominant product in every case was the reduced species **110** and, in many cases, conversion was low, with significant quantities of initial diazosulfonium conjugate **86** remaining. As a result, benzylation reaction using **128** was not pursued further.

Another factor which was investigated during the development of the photoredox benzylation reaction was the solubility of the C4-substituted Hantzsch ester reagents. It was observed that the reaction mixture was heterogeneous, with the substituted Hantzsch ester seemingly insoluble in the reaction solvent. As a result, an analogous Hantzsch ester **136** was synthesised by Dr M. T. Taylor which incorporated glycol ether groups to aid aqueous solubility (Scheme 63). **136** could be synthesised from the corresponding aldehyde, 2-phenylpropanal **134**, 2-methoxyethylacetoacetate **135** and ammonium acetate in 55% yield.



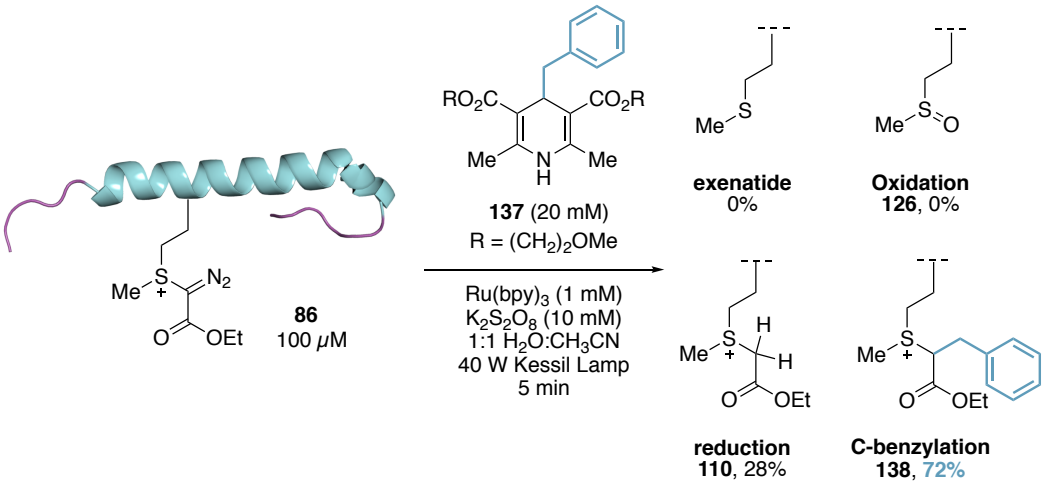
Scheme 63 Synthesis of C4-substituted Hantzsch ester incoroporating glycol groups to improve aqueous solubility. Reaction carried out by Dr M. T. Taylor.

Pleasingly, Hantzsch ester **136** displayed improved aqueous solubility and was hence employed by Dr M. T. Taylor under the photoredox alkylation conditions. A number of different conditions were screened, altering the photocatalyst loading and the amount of oxidant, improving slightly on the previous result of 38% conversion (Entry 1, Table 4). The highest conversion achieved was through lowering the concentration of Hantzsch ester **136** to 20 mM and using 1 mM photocatalyst (Scheme 64). In this case, the main competing reaction pathway was now cleavage of the methionine label, regenerating free exenatide in 23% conversion. Reduction of the diazo moiety to the methylene unit was still observed as a competing pathway, forming reduction product **110** in 10% conversion, although no oxidised peptide **126** was observed.



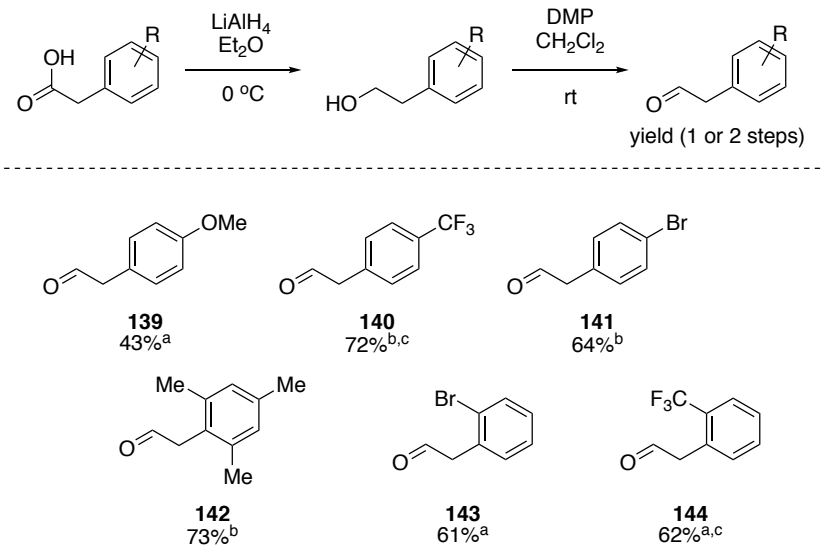
Scheme 64 Improved conditions for alkylation of exenatide conjugate using Hantzsch ester **136** with improved solubility

Following this positive result, improving the solubility of the C4-substituted Hantzsch ester reagent, the nature of the benzyl radical was modified to investigate the effect of both steric and electronic parameters. Benzyl Hantzsch ester **137** was synthesised by Dr M. T. Taylor and was subjected to visible light irradiation with sulfonium conjugate **86** in the presence of a photocatalyst and oxidant. Pleasingly, 72% conversion to benzylated product **138** was achieved, with 28% of the reduced product **110** also observed (Scheme 65). With this promising hit, it was supposed that variation of the electronic properties of the benzyl ring may affect the reactivity.



Scheme 65 Reaction of benzyl Hantzsch ester **137** under photoredox alkylation conditions, yielding C-benzylation product **138** in a promising 72% conversion. Reaction carried out by Dr M. T. Taylor.

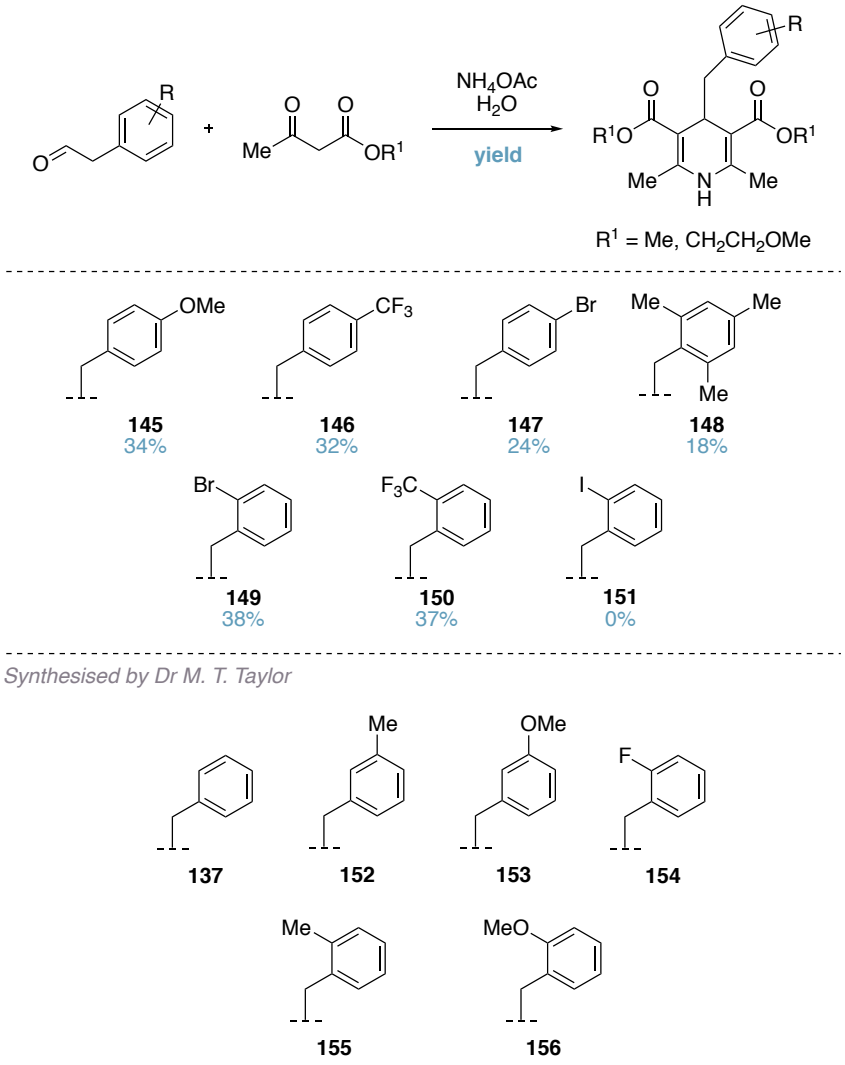
To this end, a number of C4-substituted Hantzsch esters were synthesised, with systematic variance of the substituents on the benzyl group. Firstly, synthesis of the 2-phenylacetaldehyde derivatives was carried out from the corresponding carboxylic acid, through a two-step $LiAlH_4$ reduction to the alcohol and subsequent oxidation to the aldehyde using Dess-Martin Periodinane (DMP). For some substrates, the aldehyde could be synthesised directly from the commercially available alcohol, using DMP oxidation (Scheme 66).



Scheme 66 Synthesis of a variety of aldehydes, precursors to benzyl substituted Hantzsch esters. [a]Synthesised directly from alcohol, [b]Synthesised from acid, [c]Due to lower purity aldehyde used directly in following step.

With a number of aldehydes in hand, various Hantzsch esters could be synthesised, albeit in generally poor yields (**145-150**, 18-38%, Scheme 67). In a similar manner to the tertiary and quaternary substituted Hantzsch esters, the benzyl Hantzsch esters were isolated in low yields

due to the instability of the products to dealkylation and aromatisation. As such, the oxidised Hantzsch pyridine was also isolated in significant quantities from the reactions. Unfortunately, iodo-substituted Hantzsch ester **151** could not be isolated as purification proved incredibly difficult and predominantly Hantzsch pyridine was isolated.



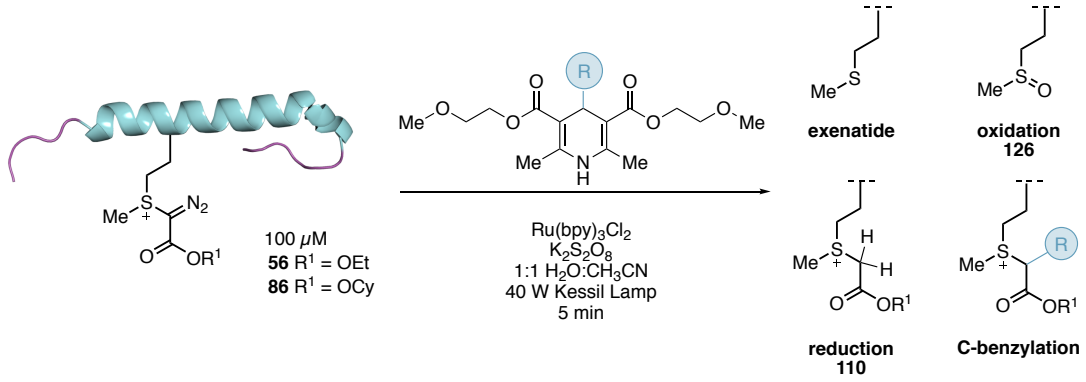
Scheme 67 Synthesis of various benzyl substituted Hantzsch esters bearing glycol groups for improved aqueous solubility.

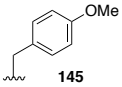
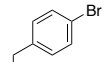
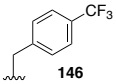
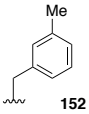
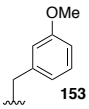
With a range of benzyl substituted Hantzsch esters in hand, incorporating both electron withdrawing and electron donating substituents, photoredox conditions could be screened in the hope of improving the conversion to C-benzylation product. Each Hantzsch ester was screened under multiple reaction conditions, varying both the concentration of Hantzsch ester and photocatalyst loading. The aim of the screening was to reduce the amount of competing reduction product, **110**, formed and hence gain a high ratio of benzylation:reduction.

It was observed that *para*-substituted benzyl Hantzsch esters gave poor conversion to product (Entries 3-6, Table 5), with neither electron donating or electron withdrawing substituents

tolerated in the reaction. *para*-Bromo substituted Hantzsch ester **147** achieved up to 20% conversion to product, though undesired formation of the reduction product **110** was the predominant side reaction (Entries 4 & 5, Table 5). *meta*-Substituted Hantzsch esters, tested by Dr M. T. Taylor (Entries 7 & 8, Table 5), were not tolerated in the reaction, resulting in complete decomposition of the polypeptide scaffold.

Table 5 Optimisation of photochemical alkylation of sulfonium diazo conjugate through systematic variance of the substituents on the benzyl transferring group. Reactions carried out by Dr M. T. Taylor are denoted in the table.



Entry	R	[Hantzsch]	[Ru(bpy)$_3$]	[$K_2S_2O_8$]	%conversion	% C-benzyl[n]	% Free peptide	% 126	% 110	Product/110 ratio
1[b,c,f]	 137	20 mM	1 mM	10 mM	>95	72	trace	0	28	**2.6/1**
2[b,c,f]		10 mM	500 μM	10 mM	36	45	12	8	36	**1.3/1**
3[b,d]	 145	40 mM	1 mM	10 mM	>95%	0	12	0	88*	-
4[b,d]	 147	20 mM	1 mM	10 mM	>95%	13	13	6	68	**0.2/1**
5[b,d]		30 mM	1 mM	10 mM	>95%	20	18	0	62*	**0.3/1**
6[b,d]	 146	40 mM	500 μM	10 mM	92%	8	5	0	79	**0.1/1**
7[b,d,f]	 152	20 mM	1 mM	10 mM	>95%	Decomposition				
8[b,d,f]	 153	40 mM	500 μM	10 mM	>95%	Decomposition				

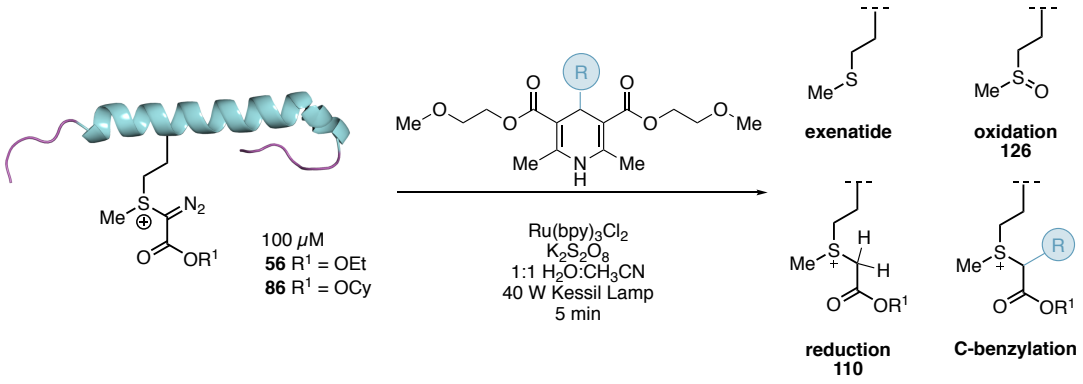
[a]Light placed at 5 cm distance from reaction vessel. [b]Reaction outcome determined by TIC. [c]ethyl ester. [d]cyclohexyl ester [e] Average of 5 runs * Reduction includes some reduction + ox [f] Reaction carried out by Dr. Michael Taylor

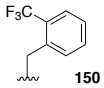
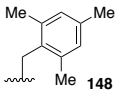
ortho-Substituted Hantzsch esters were next investigated. An electron withdrawing trifluoromethyl substituent in Hantzsch ester **150** gave poor conversion to benzylation product, despite a reasonable ratio of benzylation to reduction of 3:1 (Entry 1, Table 6). Reaction with the *ortho*-fluoro-substituted Hantzsch ester **154** proved capricious, with results ranging from complete decomposition to 45% conversion to product (Entries 2 & 3, Table 6). *ortho*-bromo Hantzsch ester **149** gave poor conversion to benzylation product, with the reduction product **110** predominating (Entry 4, Table 6). The *ortho*-methoxy substituted Hantzsch ester **156** gave reasonable conversion to product in the range of 51-54% (Entries 5 & 6, Table 6) but demonstrated a poor benzylation to reduction ratio of only 1.5:1. Given the increased reactivity observed for the *ortho*-methoxy and *ortho*-trifluoromethyl Hantzsch esters compared to their *para*-substituted analogues, it was deduced that steric effects were more strongly influencing over the benzylation reaction than electronic effects.

ortho-Methyl and mesityl Hantzsch esters **155** and **148** were observed as the most effective for benzylation (Entries 7-10, Table 6). Through increasing the concentration of Hantzsch ester from 20 mM to 30 mM, mesityl Hantzsch ester **148** gave good conversion to benzylation product, with a benzylation to reduction ratio of 8:1 (Entry 8, Table 6). Under the same reaction conditions, *ortho*-methyl Hantzsch ester **155** gave an excellent 86% conversion to product, with a benzylation to reduction ratio of 12:1 (Entry 10, Table 6), providing the best result from this screening.

It was reasoned that increased steric hindrance for the *ortho*-methyl and mesityl Hantzsch esters **155** and **148** formed a more stable, persistent radical which precluded deleterious side-reactions such as homocoupling of the resultant benzyl radical. It may be the case that the steric hindrance of the mesityl Hantzsch ester was too large, slightly inhibiting the desired radical-radical coupling, providing an optimal result for the slightly less hindered *ortho*-methyl Hantzsch ester.

Table 6 Optimisation of photochemical alkylation of sulfonium diazo conjugate through systematic variance of the substituents on the benzyl transferring group. Reactions carried out by Dr M. T. Taylor are denoted in the table.



Entry	R	[Hantzsch]	[Ru(bpy)$_3$]	[$K_2S_2O_8$]	%conversion	% C-benzyl[n]	% Free peptide	% 126	% 110	Product/110 ratio
1[b,d]	 150	30 mM	1 mM	10 mM	53%	32	8	3	11	**3.0/1**
2[b,d,f]	 154	20 mM	1 mM	10 mM	>95%	Decomposition				
3[b,d,f]		10 mM	500 μM	5 mM	67%	45	6	trace	16	**2.9/1**
4[b,d]	 149	20 mM	1 mM	10 mM	71%	37	10	4	16	**0.8/1**
5[b,d,f]	 156	20 mM	1 mM	10 mM	>95%	54	5	4	36	**1.5/1**
6[b,d,f]		10 mM	500 μM	5 mM	>95%	51	trace	4	43	**1.2/1**
7[b,d]	 148	20 mM	1 mM	10 mM	>95%	70	0	0	22	**3.2/1**
8[b,d]		30 mM	1 mM	10 mM	89%	72	6	0	9	**8.0/1**
9[b,d,f]	 155	20 mM	1 mM	10 mM	>95%	72	8	trace	19	**3.2/1**
10[b,d,e,f]		30 mM	1 mM	10 mM	>95%	86	3	trace	7	**12/1**

[a]Light placed at 5 cm distance from reaction vessel. [b]Reaction outcome determined by TIC. [c]ethyl ester. [d]cyclohexyl ester [e] Average of 5 runs * Reduction includes some reduction + ox [f] Reaction carried out by Dr. Michael Taylor

With optimised conditions for the benzylation of the diazosulfonium species in hand, it was proposed that a derivative of the C4-substituted Hantzsch ester could be synthesised bearing a functional payload. If successfully implemented in the photoredox benzylation reaction, the transfer of a biologically relevant functional group would be achieved, using the diazosulfonium species as a platform for bioorthogonal functionalisation. The small-molecule probe biotin is commonly used for protein detection and purification due to its extremely strong

binding to the proteins avidin and streptavidin.[252] As a result, the introduction of biotin probes to biomolecules, using bioorthogonal or bioconjugation techniques, is relatively routine.[4] It was proposed that the introduction of biotin using this photoredox benzylation would provide a useful proof-of-principle result for the transfer of biologically relevant functionality.

A C4-substituted Hantzsch ester was designed bearing a biotin probe (**157**, Figure 14). Key structural features from the benzyl Hantzsch ester optimisation were incorporated in the hope of maintaining the desired reactivity, namely, the glycol ester groups for improved aqueous solubility and the *ortho*-methyl group on the benzyl substituent. As it had been established that methyl substituents on the benzyl ring were tolerated, an alkyl linker to the payload was also designed. A retrosynthetic analysis of **157** was proposed, incorporating the biotin payload as the final step of the synthesis from key intermediate **158** using activated ester **159**. **158** could be synthesised in the same manner as the benzyl C4-substituted Hantzsch esters, from the corresponding aldehyde, **160**, 2-methoxyethyl acetoacetate and ammonium acetate.

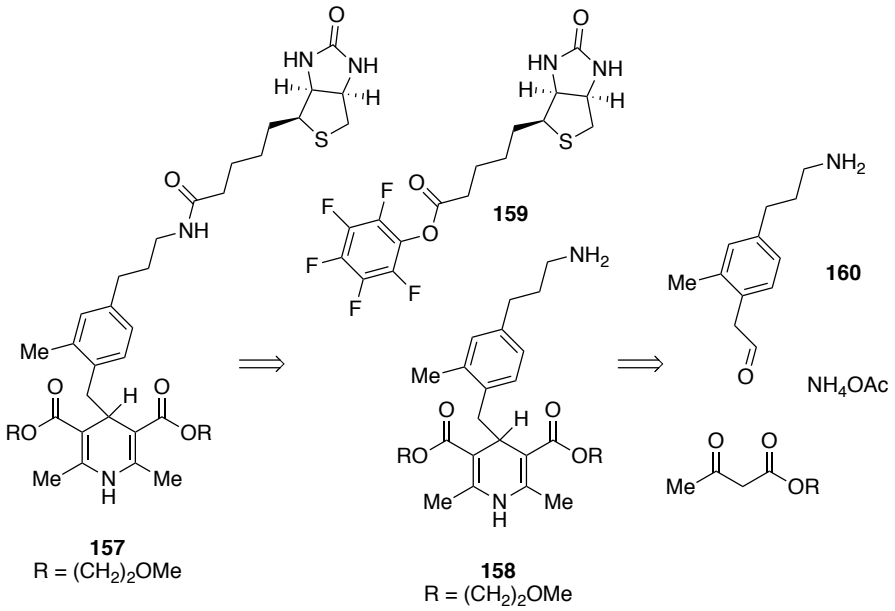
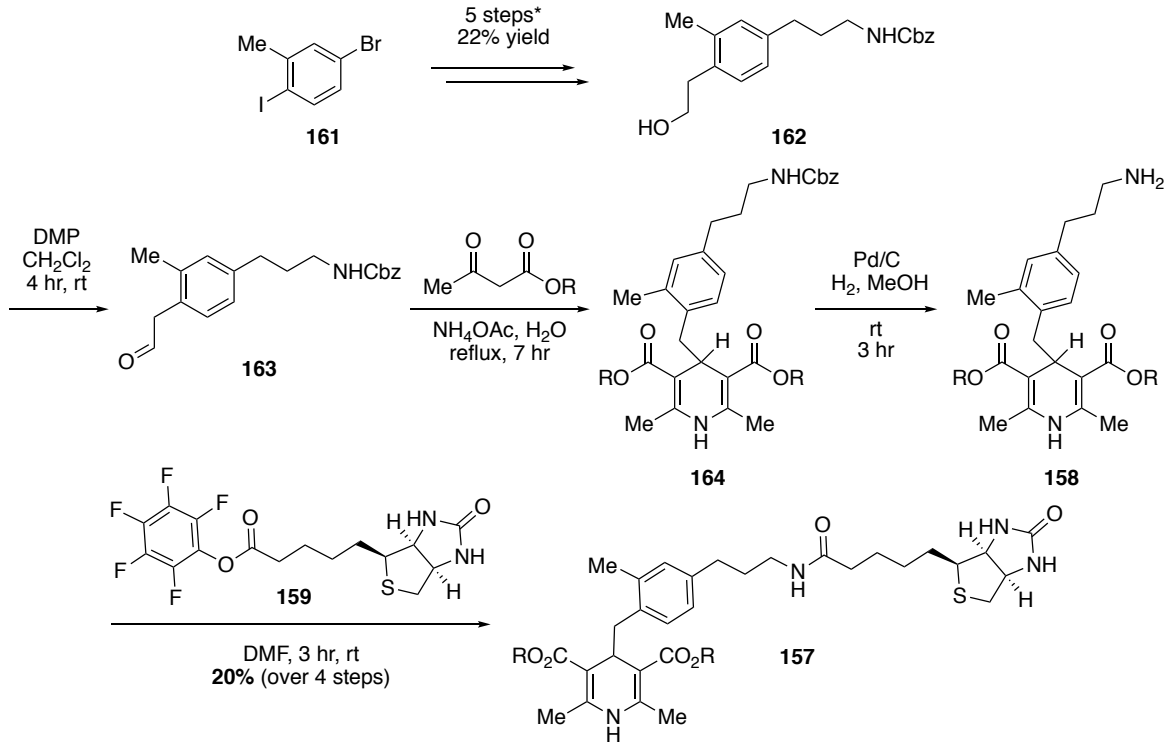


Figure 14 Design of Hanztsch ester bearing functional payload (biotin) and retrosynthetic analysis

The synthesis of Hantzsch ester **157** bearing biotin as a functional payload was synthesised in collaboration with Dr C. He from commercially available 5-bromo-2-iodotoluene **161** (Scheme 68). Alcohol **162** was synthesised in 22% yield over 5 steps. **162** could then be oxidised to the corresponding aldehyde **163** using Dess-Martin periodinane, yielding the key aldehyde intermediate. Aldehyde **163** was subjected to the standard conditions used previously for the synthesis of the Hantzsch ester scaffold, followed by hydrogenation to remove the amine

protecting group, affording **158**. Final incorporation of the biotin payload was achieved through reaction with activated ester **159** to deliver the final compound, **157**, in 20% yield over 4 steps.

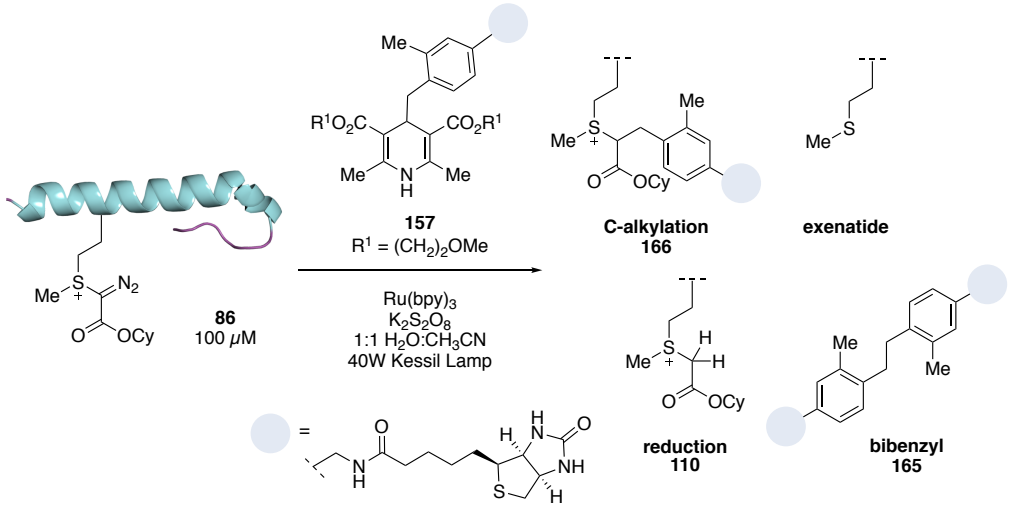


Scheme 68 Synthesis of biotin Hantzsch ester **157**. *Carried out by Dr C. He.

The Hantzsch ester bearing the biotin functional handle **157** was subsequently screened in the benzylation reaction with exenatide sulfonium conjugate **86**. In all cases, the reaction profile was complex, preventing quantitative determination of resulting products. Therefore, a qualitative determination of the reaction outcome was recorded by LCMS, determining the major and minor products observed. Unfortunately, transfer of the benzyl radical with appended biotin payload could not be observed. The predominant product observed under all conditions screened was the reduced species **110** (Entries 1-3, Table 7). Interestingly, under the standard conditions screened in the reaction, the bibenzyl species **165** was observed in all cases. This observation indicates the likely formation of the benzyl radical and suggests that radical formation was not the limiting factor in the formation of alkylation product **166**. Control reactions (Entries 4-6, Table 7) showed that in the absence of the oxidant the bibenzyl species could be observed, but only on increasing the reaction time to 20 minutes. Interestingly, the bibenzyl species could also be observed in the absence of the photocatalyst. As the Hantzsch ester bears a hydrogen atom as well as the benzyl group at the 4-position, it is possible that

hydrogen atom transfer is much more facile compared to transfer of the larger, biotin-derived benzyl radical, causing the reduction product to form preferentially.

Table 7 Attempted alkylation of diazosulfonium conjugate **86** using Hanztsch ester **157**. Predominant product formed was reduction product **110**. Due to messy reaction profile, outcomes of reactions were determined in a qualitative manner.

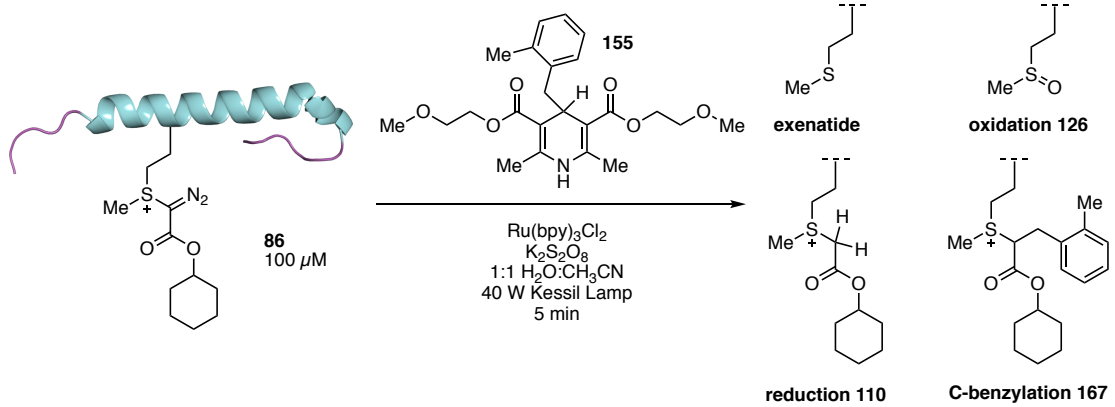


Entry	[Hantzsch]	[Photocat.]	[$K_2S_2O_8$]	T(min)	86 remaining	reduction 110	C-benzyl^n 166	exenatide	bibenzyl 165 observed?
1[a,b]	30 mM	1 mM	10 mM	5	—	—	—	—	Yes
2[a,b]	20 mM	1 mM	1 mM	5	—	major	—	trace	Yes
3[a,b]	20 mM	1 mM	0.5 mM	5	—	major	—	minor	Yes
4[a,b]	20 mM	1 mM	—	5	—	major	—	—	No
5[a,b]	30 mM	1 mM	—	20	—	major	—	—	Yes
6[a,b]	20 mM	—	10 mM	5	major	—	—	—	Yes

[a)]Light placed at 5 cm distance from reaction vessel. [b)]Reaction outcome determined by TIC.

It was hoped that by carrying out additional control reactions on the simple benzylation reaction, it would be possible to gain some insight into the role of each reagent. It was found that in the absence of the persulfate oxidant, the conversion was dramatically reduced, with 44% of starting diazo conjugate remaining after the reaction and only 4% of benzylation product **167** observed (Entry 2, Table 8). The exclusion of the photocatalyst, $Ru(bpy)_3Cl_2$, impeded the reaction further, with only 25% conversion of starting material and only trace levels of benzylation product observed (Entry 3). Finally, carrying out the reaction in the absence of visible light resulted in exclusive decomposition of the polypeptide (Entry 4). These relevant control reactions indicated the necessity of each component for benzylation. Taken together, these results suggest a visible-light mediated photoredox process.

Table 8 Control experiments run to investigate the role of each reagent. Reactions carried out by Dr M. T. Taylor are denoted

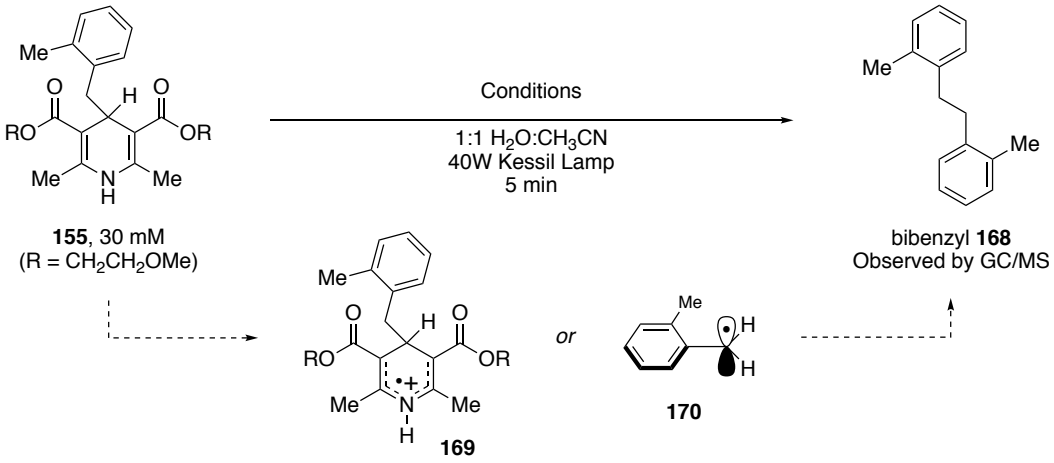


Entry	[Hantzsch]	[Ru(bpy)$_3$]	[$K_2S_2O_8$]	Light[a]	%conversion[b]	% Free peptide	% 126	% 110	% C-benzyl[n] 167
1[c]	30 mM	1 mM	10 mM	40W kessil	>95	3	trace	7	**86**
2[e]	30 mM	1 mM	—	40W kessil	56	3	trace	44	**4**
3[e]	30 mM	—	10 mM	40W Kessil	25	3	7	14	trace
4[e]	30 mM	1 mM	10 mM	dark[d]	Decomposition				

[a)]Light placed at 5 cm distance from reaction vessel. [b)]Reaction outcome determined by TIC. [c)] Values are an average of 5 experiments. [d)]reaction was run in a foil-encapsulated vial. [e)]Reaction carried out by Dr M. T. Taylor.

Additional control reactions were carried out by Dr M. T. Taylor, attempting to observe the formation of bibenzyl species **168** (Table 9). It was proposed that the formation of this bibenzyl species may be indicative of the formation of radical species **169** or **170**, which may give insight into the mechanism of the benzylation reaction. Interestingly, **168** was only observed in the presence of both the photocatalyst and oxidant, eluding to the importance of both of these reagents in the formation of the benzyl radical alkylating species.

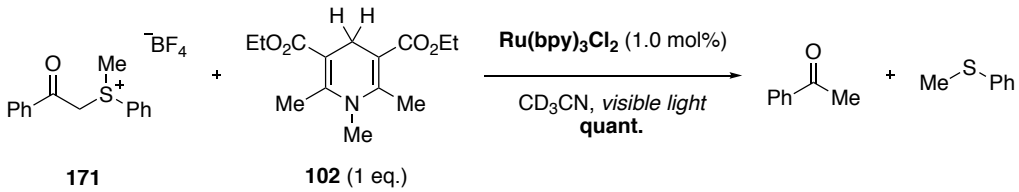
Table 9 Control reactions to assist in the identification of the role of each reagent. Conditions which formed bibenzyl **168** would suggest the formation of either of two radical species **169** or **170**. Reactions carried out by Dr M. T. Taylor.



Entry	$[Ru(bpy)_3]$	$[K_2S_2O_8]$	Light[a]	168 Observed?
1	1 mM	10 mM	40W kessil	Yes
2	1 mM	—	40W kessil	No
3	—	—	40W Kessil	No

[a)]Light placed at 5 cm distance from reaction vessel

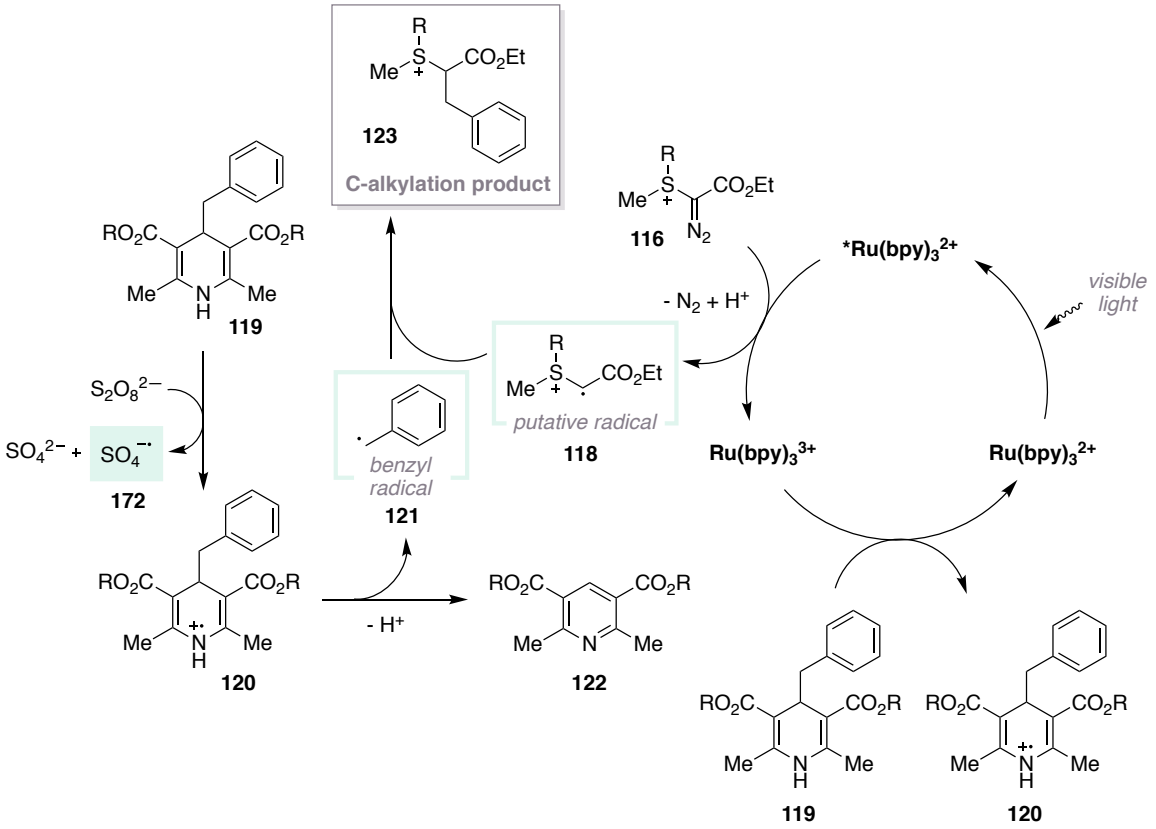
One of the earliest reported photoredox-catalysed transformations of organic molecules was reported by Kellogg in 1978 for the reduction of phenacyl sulfonium salts by 1,4-dihydropyridines **102** (Scheme 69).[246,253] In the presence of visible light, the reduction of sulfonium salts **171** was observed to be dramatically accelerated in the presence of various photosensitisers, with $Ru(bpy)_3Cl_2$ providing the greatest rate enhancement. Considering these early examples of sulfonium reduction, it is remarkable that no reduction of the sulfonium species generated in the methionine bioconjugation was observed on subjecting these species to similar photoredox conditions.



Scheme 69 Early photoredox-catalysed transformation reported by Kellogg for the reduction of phenacyl sulfonium salts

Using the information gained from the control reactions, a mechanism could be suggested for the formation of alkylation product **167** (Scheme 70). Upon visible light irradiation, excitation of the photocatalyst can occur, forming excited species $*Ru(bpy)_3^{2+}$. This species is highly

reducing and may donate an electron to diazosulfonium species **116**, forming the putative radical **118** through loss of N_2 and protonation.[237,238] The strongly oxidising persulfate species may facilitate the generation of the radical cation of the Hanztsch ester **120**, through single electron oxidation of the Hanztsch ester **119**, generating a sulfate anion and a sulfate radical anion **172**.[244] This is supported by the observation of dibenzyl formation in the absence of the photocatalyst, suggesting it is the oxidant and not the photocatalyst responsible for the oxidation of the Hantzsch ester. The benzyl radical **121** could be generated from the aromatisation of the radical cation **120**, forming Hantzsch pyridine **122**.[248] Subsequent radical-radical coupling of benzyl radical **121** and putative radical **118** could form the C-benzylation product **123**. Benzyl Hantzsch ester **119**, which is present in excess, may be responsible for the turnover of the photocatalyst, reducing $Ru(bpy)_3^{3+}$ back to $Ru(bpy)_3^{2+}$. Further mechanistic investigation is required to validate the proposed mechanism.



Scheme 70 Proposed mechanism for formation of C-alkylation product

2.4.7 Summary

This chapter has described the optimisation, development and characterisation of a novel, biocompatible, alkylative bioconjugation strategy targeting the amino acid methionine.[254] Initial work by Dr M. G. Suero and Dr M. T. Taylor identified this interesting reactivity, wherein hypervalent iodonium salts bearing a diazo ester moiety could selectively react at the sulfide side chain of methionine within small peptide substrates. Subsequent optimisation was carried out and identified a number of larger polypeptide and protein substrates, bearing multiple nucleophilic amino acid residues, which could be functionalised with exquisite selectivity, forming β-sulfonium α-diazo ester conjugates. The stability of the resulting functionalised species was investigated, with the half-life proving to be within the range of 30-96 hours, similar to that of established protein conjugates.[230,231] Mass spectral characteristics were examined in order to identify the nature of an observed fragmentation pattern, as well as extensive characterisation of sulfonium products through MS/MS and CD analysis. A Staudinger-inspired phosphine-mediated cleavage of the sulfonium label was developed, enabling the triggered regeneration of free methionine-containing polypeptides and proteins. A photochemical reduction, initially discovered by Dr M. T. Taylor, reduced the β-sulfonium α-diazo ester conjugates to form trialkylsulfonium species in quantitative conversion. On investigation of the half-life of the resulting reduced species, the trialkylsulfonium conjugates exhibited significantly improved stability in relation to the diazo analogue. Finally, an alkylative functionalisation was developed, using photoredox catalysis, targeting the diazo moiety formed through the initial bioconjugation reaction. Using C4-substituted Hantzsch esters, a ruthenium photocatalyst and an oxidant, functionalised benzyl groups could be transferred to the α-position of the sulfonium conjugate, providing a method for bioorthogonal functionalisation of the non-native diazo group.

3 Synthesis of multi-functionalised polypeptides and proteins

3.1 Introduction

In recent years, the field of protein modification has advanced beyond mono-functionalisation, with increasing interest in the development of dual modification strategies for increasing the exogenous functionality of peptides and proteins.[255] Interest in the area of the dual-functionalisation of proteins has grown, as the introduction of at least two functional groups is valuable for many applications; namely, targeted biotherapeutics,[256–258] Forster Resonance Energy Transfer (FRET) for monitoring of protein dynamics[259] and also protein folding studies.[260] Chemical methods for the site-selective dual-modification of proteins remain a challenge.[255] Therefore, development of operationally simple, versatile methods for the introduction of two or more modalities to a protein is fundamental to the advancement of many biologically relevant applications.

A common approach for dual-functionalisation has involved the sequential application of two established methods for protein modification, to form multi-functional scaffolds.[261] This approach generally exploits two different natural or unnatural amino acid residues (Figure 15a). A second, distinct approach involves the conjugation of a bifunctional species to a protein, bearing two orthogonally reactive functional handles. This bifunctional handle may be reacted sequentially to form dual-functionalised scaffolds[262] (Figure 15b). The latter approach has had limited success due to the complex synthetic challenge presented by the branched structures bearing orthogonal handles.[255] Additionally, this approach has rarely been exemplified on protein scaffolds. Instead, small molecules or simple peptides are more commonly exploited.[255] As a result, this chapter will focus on the synthesis of dual-functionalised proteins *via* the labelling of two different amino acid residues. For a more in-depth account of dual functionalisation using multifunctional scaffolds the reader is directed to a recent review.[255]

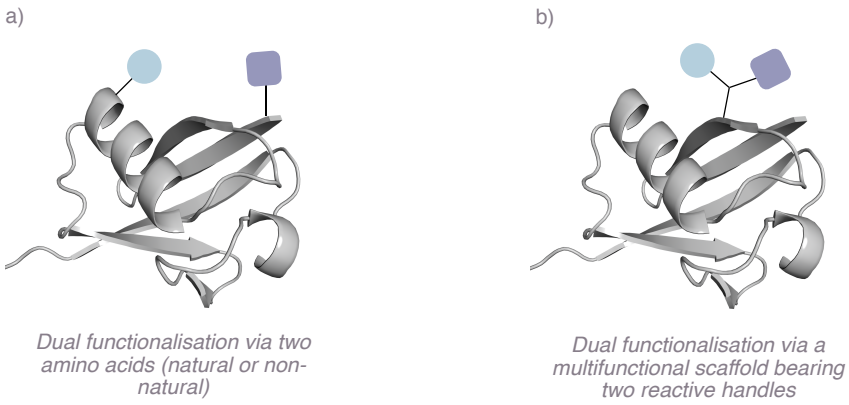
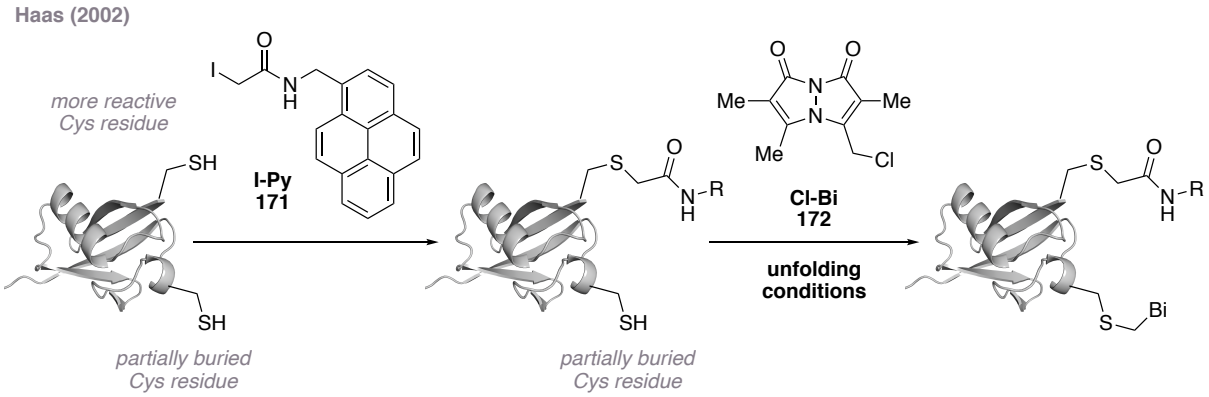


Figure 15 Two different strategies for dual functionalisation of proteins; a) Through introduction of a single multifunctional scaffold bearing two reactive handles b) Exploiting two different amino acids (natural or non-natural)[255]

3.1.1 Double labelling using identical amino acids with different reactivity

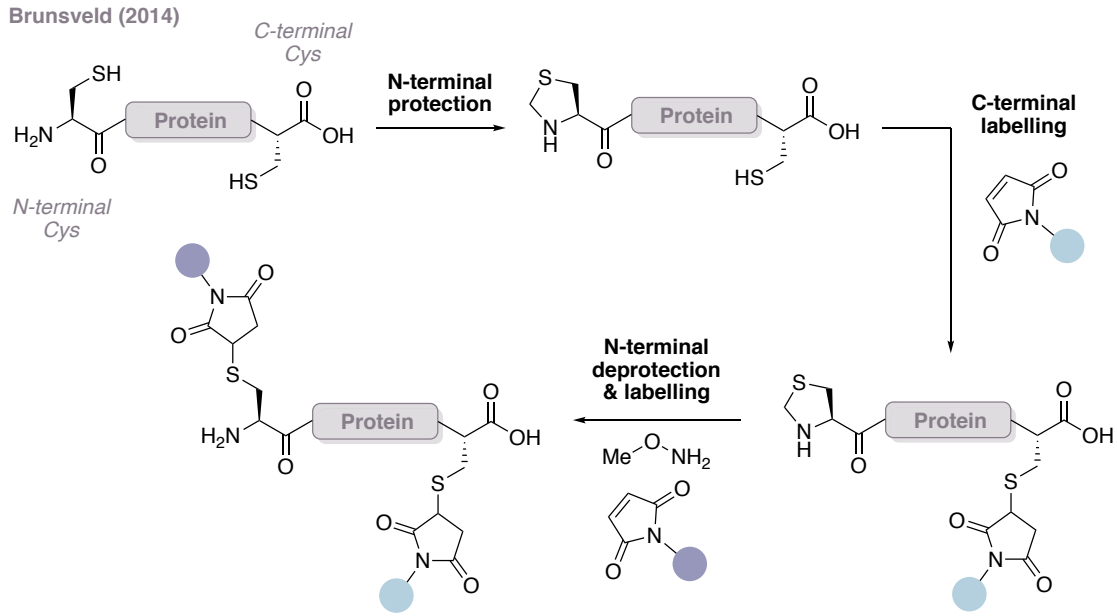
Site-selective single-functionalisation of proteins can be achieved through exploitation of differences in reactivity or accessibility of two identical amino acids.[255] It has been demonstrated that, in the case of lysine, the pKa of the amine side chain can vary by up to 5 units depending on the polarity of its microenvironment, its solvent accessibility, protein folding and the nature of the surrounding residues (Section 1.2.3).[131] Selective reaction of 'hyper-reactive' lysine residues, exploiting differences in reactivity, allows preferential single-labelling with specifically designed reagents under pH control. This concept can be applied to the dual-labelling of proteins, exploiting the differences in reactivity of other residues, including cysteine.

Haas and co-workers engineered a mutant form of *Escherichia coli* adenylate kinase, in which two cysteine residues were carefully positioned so as to induce distinct reactivity.[263] By exploiting the nucleophilicity difference between two free thiol motifs, initial reaction with iodoacetamide **171** occurred selectively at the more reactive site. Subsequently, using unfolding conditions to access the less reactive cysteine residue, a second fluorescent probe **172** could be attached selectively, forming a dually-functionalised protein scaffold which could be used in FRET measurements (Scheme 71). However, the requirement for partial denaturation of the protein by unfolding could potentially impair the protein function.



Scheme 71 Technique for the dual labelling of *Escherichia coli* mutant containing two distinct cysteine residues

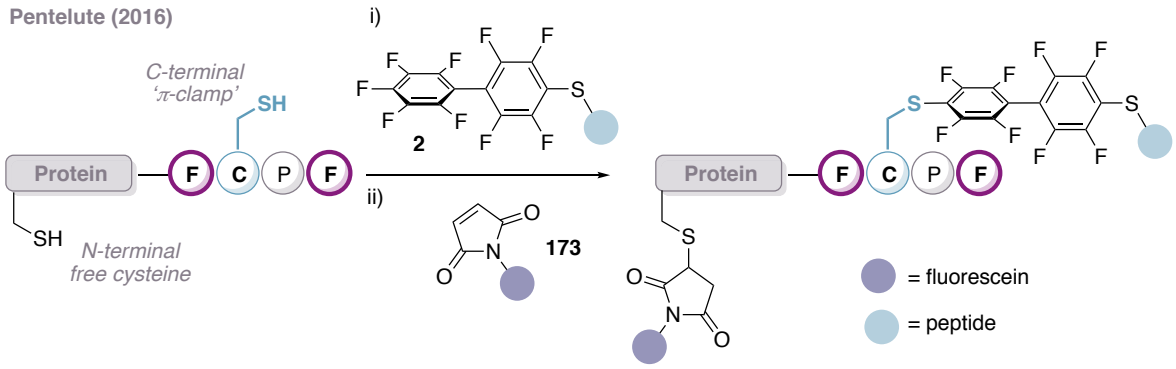
Brunsveld *et al.* described a cysteine dual labelling strategy *via* a temporary protection approach (Scheme 72).[264] Initial protection of both the N-terminus and the N-terminal cysteine side chain as a thioazolidine ring was followed by selective labelling of the C-terminal cysteine residue using a maleimide reagent. Subsequent N-terminal cysteine deprotection facilitated a second selective labelling using a different maleimide reagent. The temporary protection strategy was reported to be effective with a wide range of other site-specific ligation techniques such as native chemical ligation or disulfide exchange. However, this temporary protection and deprotection approach requires several steps to achieve high selectivity.



Scheme 72 Site-selective dual labelling with maleimide reagents through temporary protection of an N-terminal cysteine residue

In a similar strategy, detailed in Section 1.2.2, Pentelute and co-workers reported the exploitation of a four amino acid π-clamp recognition motif (Phe-Cys-Pro-Phe) for selective

perfluoroarylation of single cysteine residues within proteins.[93] The unique microenvironment of the clamp's cysteine residue was exploited in order to selectively form dually functionalised proteins. First, nucleophilic aromatic substitution was employed selectively at the π-clamp recognition unit, followed by Michael addition of an N-terminal cysteine residue to a fluorescein-derived maleimide reagent **173** (Scheme 73). Selective dual labelling in this case could be carried out in a one-pot procedure but required pre-installation of the π-clamp sequence at the C-terminus of the desired protein.

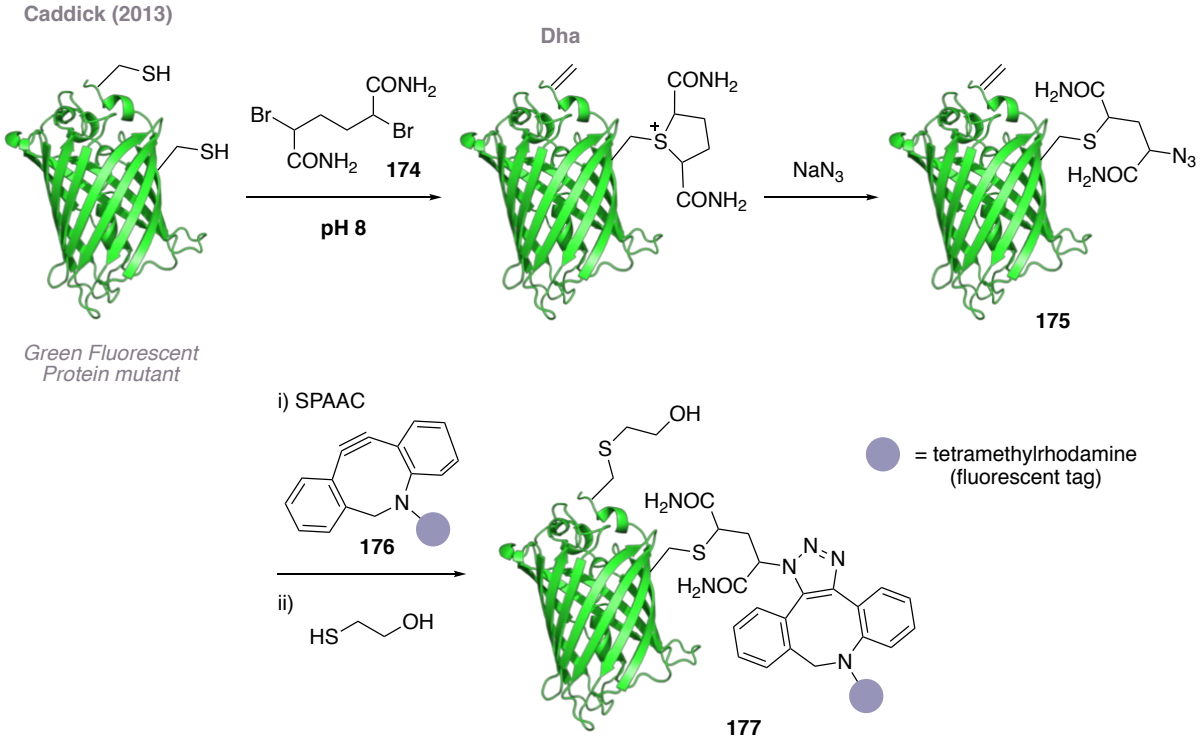


Scheme 73 Dual labelling of proteins using a pre-installed "π-clamp" recognition motif and subsequent maleimide functionalisation

These strategies, which target two cysteine residues, rely on the differing nucleophilicity of the thiol side chains and require careful incorporation of the cysteine residues at specific sites within the protein sequence. Rigorous control of the conditions, reagents and reactions times must be adhered to in order to maintain high selectivity. A distinct approach employs substrate control to achieve selective dual labelling using a single reagent, inducing differing reactivity in two cysteine residues.

Caddick and co-workers were able to exploit the natural microenvironment of Green Fluorescent Protein (GFP). They showed that a reaction between two cysteine residues and 2,5-dibromohexanediamide **174** could result in a dual modified protein (Scheme 74).[265] Initial reaction of a double cysteine mutant of GFP with **174** formed a bis-sulfonium species. The system was engineered so that one cysteine residue had an accessible α-proton, and therefore readily eliminated a neutral sulfide species to form dehydroalanine (Dha). The α-proton of the second cysteine residue was sufficiently shielded, and it therefore persisted as a stable sulfonium moiety. Chemoselective ring opening of the sulfonium motif with sodium azide furnished **175**, bearing two orthogonal reactive handles. Sequential strain-promoted azide alkyne cycloaddition with dibenzylcyclooctyne reagent **176**, incorporating a fluorescent dye,

was followed by Michael addition of 2-mercaptoethanol to Dha, yielding dual labelled GFP, **177**.

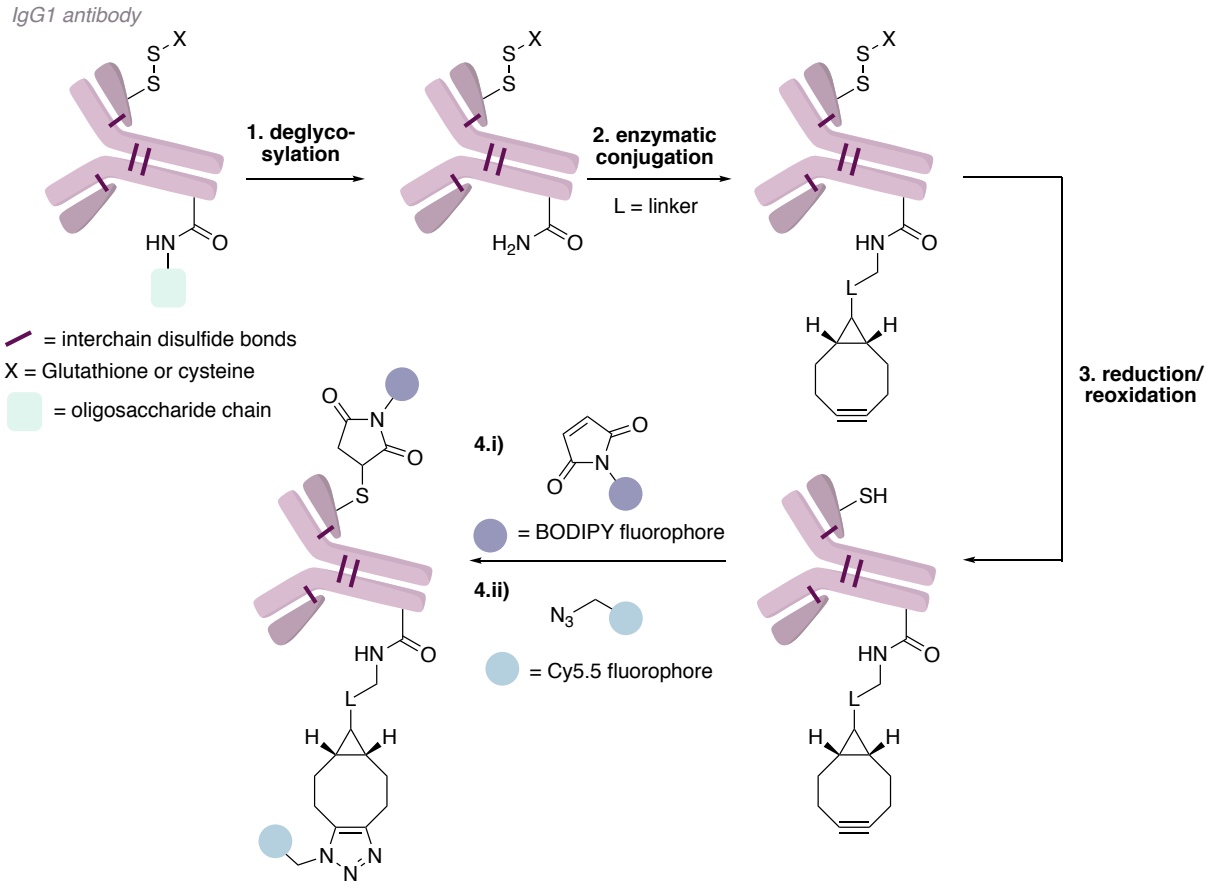


Scheme 74 Exploitation of the microenvironment of Green Fluorescent Protein for distinct reactivity with reagent **174** to introduce two different orthogonal handles

3.1.2 Double labelling using two different amino acids

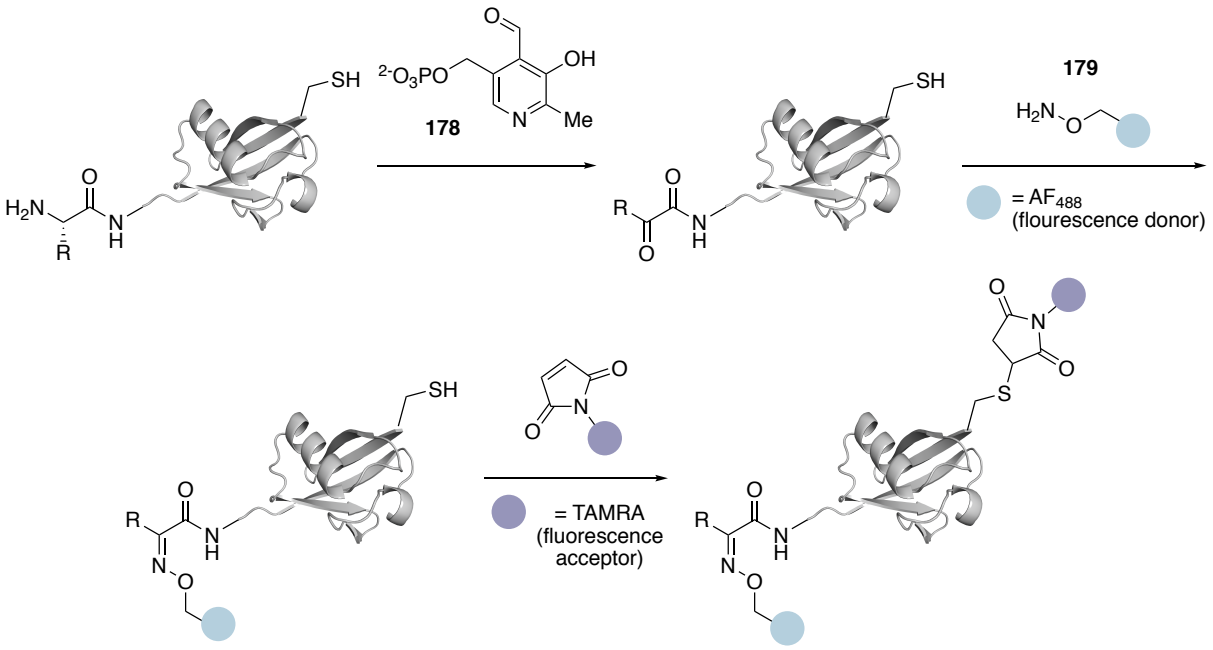
Rather than exploiting two of the same amino acid residues that exhibit different reactivity, different amino acid residues with orthogonal reactivity may be targeted. Puthenveetil *et al.* used this approach to generate dual labelled antibody conjugates by exploiting two different native sites on an Immunoglobulin G1 (IgG1) antibody (Scheme 75).[266] Deglycosylation, the removal of oligosaccharide chains which are covalently attached to amino acid side chains, was followed by enzymatic conjugation to allow the selective introduction of the first reactive, orthogonal handle (Steps 1 & 2). A reduction and reoxidation sequence reduced the engineered cysteine residues on the exterior of the antibody, while retaining the interchain disulfide bonds (Step 3). Michael addition to the free thiol residue with maleimide reagents, to introduce the first fluorescent probe, was followed by strain-promoted azide-alkyne cycloaddition to introduce the second probe molecule (Step 4i & ii). This strategy was reported to be virtually compatible with all immunoglobulin G antibodies as cysteine incorporation has been shown to

be facile and all IgGs share the same glycosylation site, to which oligosaccharide chains are attached post-translationally.[255] However, the method required several steps, making dual labelling laborious.



Scheme 75 Formation of dual functionalised antibody conjugates targeting native amino acid residues

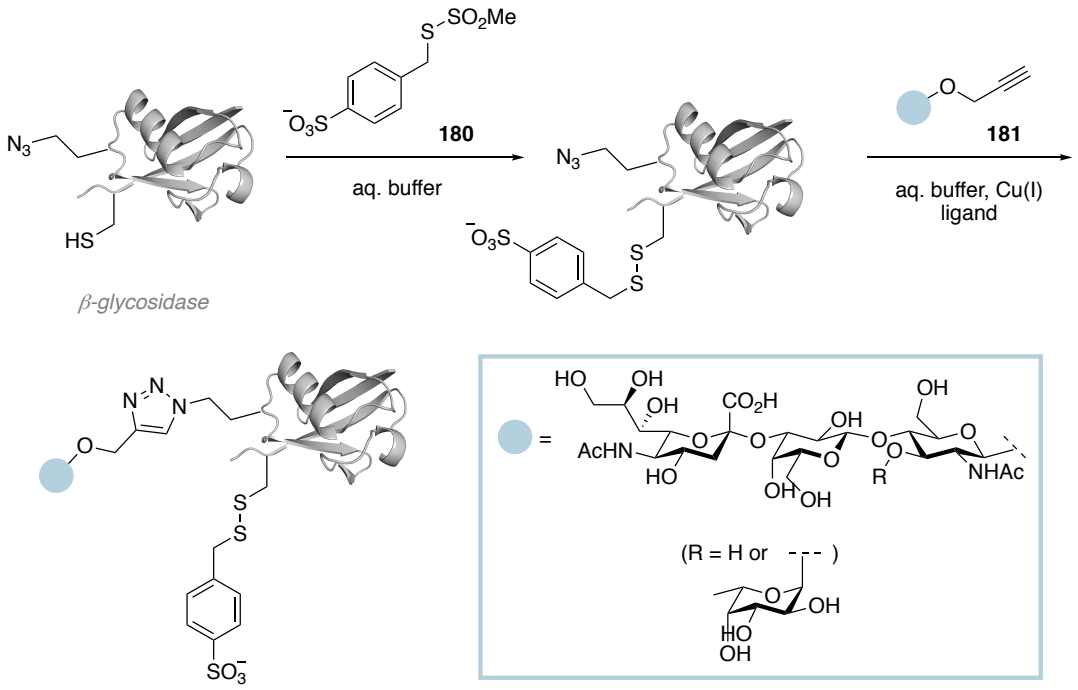
Paavola *et al.* described the combination of traditional cysteine-maleimide chemistry with a transamination reaction at the N-terminus (Scheme 76).[267] Methodology developed by Francis and co-workers was used to install a uniquely reactive ketone or aldehyde at the N-terminus of the protein with pyridoxal-5'-phosphate **178**.[148] Subsequent oxime formation was carried out with alkoxyamine reagents bearing a donor fluorescent payload **179**. Maleimide chemistry could then be employed to incorporate an acceptor fluorophore at a free cysteine residue, selectively forming dual functionalised protein species which were used in FRET analysis.



Scheme 76 Dual functionalisation of proteins targeting the N-terminus and a cysteine residue for the introduction of two complementary fluorophores

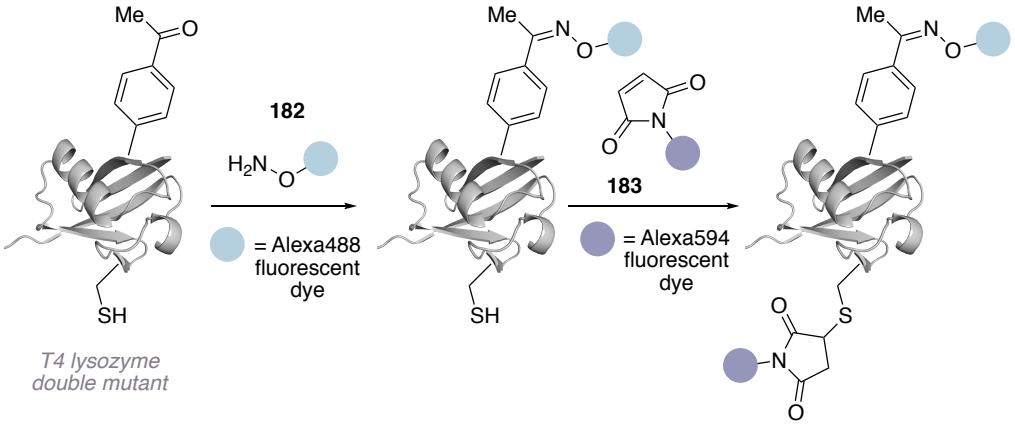
Another potential strategy for dual functionalisation of biomolecules is to target two non-native functional groups in a bioorthogonal fashion. However, synthetic genetic code expansion for the incorporation of non-natural amino acids is generally limited to the incorporation of a single type of non-natural amino acid.[268] Methods for the incorporation of multiple non-natural residues have been developed independently by Chin[268] and Liu.[269] However, this dual incorporation approach has often been associated with low protein yields.[270,271] As a result, bioorthogonal strategies for dual labelling generally involve the targeting of one non-native and one native amino acid residue.

An example of this approach was reported by Davis *et al.* (Scheme 77).[272] Incorporation of the non-natural amino acid azidohomoalanine (Aha) facilitated the use of the copper-catalysed azide-alkyne cycloaddition (CuAAC) which could be carried out with a tri- or tetrasaccharide derived alkyne reagent **181**. Disulfide ligation could be used to introduce a modification at a free cysteine residue using sulfotyrosine mimic **180**. The mutual compatibility of the two modification reactions was demonstrated by the ability to reverse the order of the sequential functionalisations. Additionally, it was shown that different chemical tags could be utilised, such as homopropargylglyine amino acid (Hpg), which underwent subsequent reaction with azide reagents. The authors demonstrated the applicability of this technique towards the detection of mammalian brain inflammation and disease.[272]



Scheme 77 Schematic depiction of the exploitation of non-natural amino acid azidohomoalanine (Aha) and cysteine in the synthesis of dual functionalised protein scaffolds

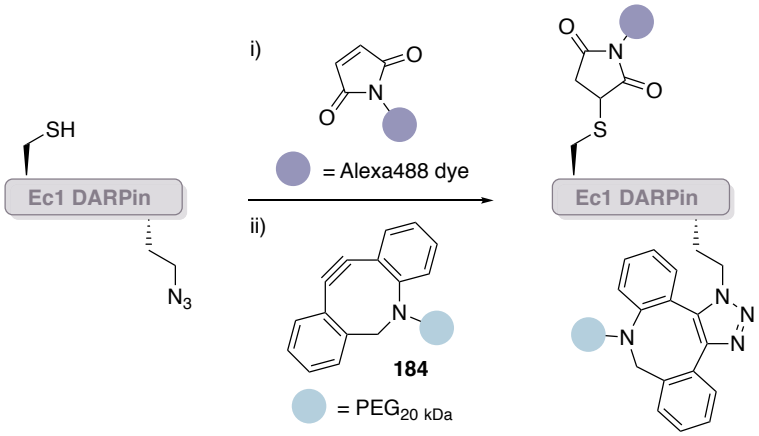
A combination of native and non-native amino acid functionalisation was also employed by Schultz and Deniz to study the structure and folding of T4 lysozyme. The non-native residue *p*-acetylphenylalanine was incorporated and functionalised with a fluorescent probe (Alexa488 dye, **182**), using a selective oxime ligation (Scheme 78).[273] Subsequent Michael addition at the free cysteine residue with a maleimide reagent **183** selectively introduced a second fluorescent tag (Alexa594). The resulting dual-labelled species could be used in FRET analysis.



Scheme 78 Dual labelling of T4 lysozyme with two different fluorescent tags through the use of non-native amino acid *p*-acetylphenylalanine and native amino acid cysteine.

In a similar strategy to that of Davis *et al.* (Scheme 77), Pluckthun and co-workers developed a dual functionalisation protocol of the non-native amino acid azidohomoalanine (Aha) and

native amino acid cysteine to form dually functionalised antibody mimetic species known as Designed Ankyrin Repeat Proteins (DARPins) (Scheme 79).[274] DARPins are a leading alternative to antibody proteins for selective binding to molecular targets. In contrast to the work described by Davis, Pluckthun opted for selective strain-promoted azide-alkyne cycloaddition with dibenzylcyclooctyne reagent **184** and maleimide modification of a single cysteine residue. The resulting functionalised DARPins could be used to measure association and dissociation kinetics on tumour cells to observe the dynamic blocking effect of the PEG chains on DARPin binding.



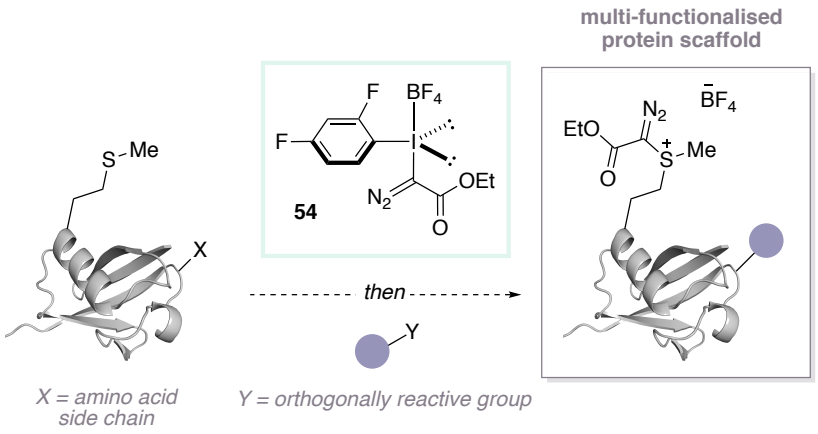
Scheme 79 Dual functionalisation of DARPins exploiting the non-native amino acid azidohomoalanine (Aha) and native amino acids cysteine

Many methods for the selective, multi-functionalisation of proteins have been developed in recent years. However, lengthy syntheses are often required in order to obtain the modified scaffolds, and cross reactivity can cause problems with selectivity.[255] As a result, the development of additional, tailored strategies for the formation of dual-functionalised proteins are required to overcome limitations with existing technologies. Alternative methods for the facile synthesis of multi-functionalised protein conjugates would provide access to species which are integral to a number of biologically relevant applications such as FRET or targeted biotherapeutics.

3.2 Project aims

Currently, there are no methods which employ methionine functionalisation as a strategy for the synthesis of multi-functionalised protein scaffolds. The use of methionine as an orthogonal handle, alongside other amino acids, in the dual functionalisation of proteins may address some limitations of current strategies, such as lengthy synthesis and variable product stability.

The aim of this project was to investigate the orthogonality of the methionine-selective bioconjugation reaction, developed in Chapter 2, with other protein functionalisation methods (Scheme 80). It was envisaged that two bioconjugation techniques could be used in combination, enabling the dual-functionalisation of native polypeptides and proteins without the need for prior incorporation of non-native groups.

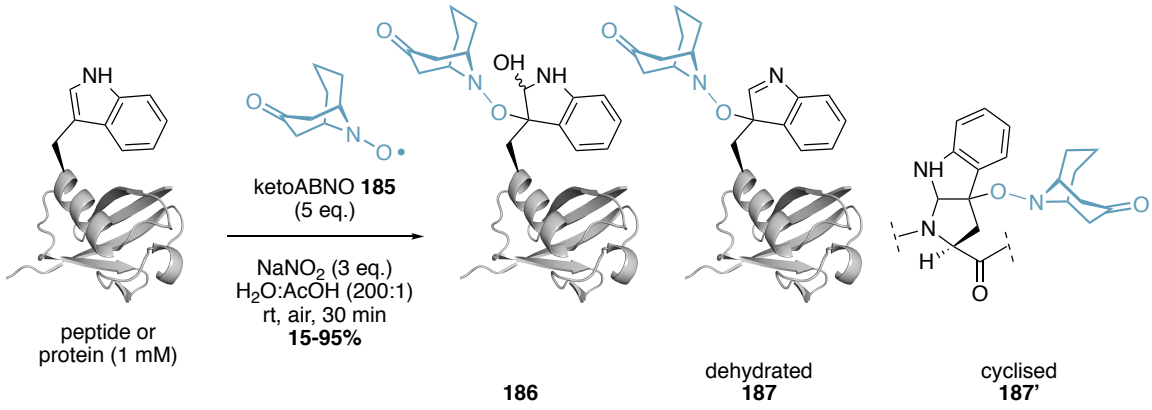


Scheme 80 Schematic representation of proposed synthesis of dual functionalised protein constructs exploiting methionine-selective bioconjugation approach with hypervalent iodine reagents

3.3 Results and discussion

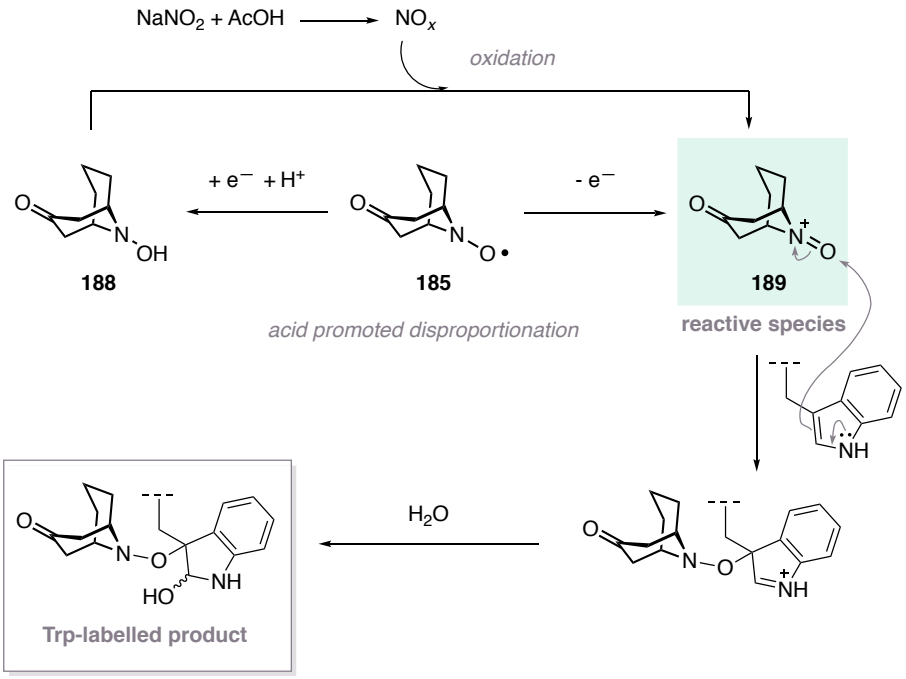
3.3.1 Dual-labelling of methionine and tryptophan residues

In 2016, Kanai and co-workers reported a transition metal-free tryptophan-selective bioconjugation.[139] Using a bicyclic organoradical species, ketoABNO **185**, and an oxidant, sodium nitrite, site-selective functionalisation of tryptophan residues within peptides and proteins could be achieved in up to 95% conversion (Scheme 81). Formation of a C–O bond at the 3-position of the indole moiety afforded product **186**, and additional products were formed through dehydration of **186** across the indole C–N bond, forming dehydrated **187**, or dehydration of the polypeptide backbone, forming cyclised **187'**. The authors reported the conversion to product as the total of the three functionalised species; **186**, **187** and **187'**, although no ratio between products was reported.



Scheme 81 Transtition metal-free tryptophan-selective bioconjugation of polypeptides and proteins. Dehydration of product **186** can occur either across the C–N bond of the indole moiety or across the protein backbone

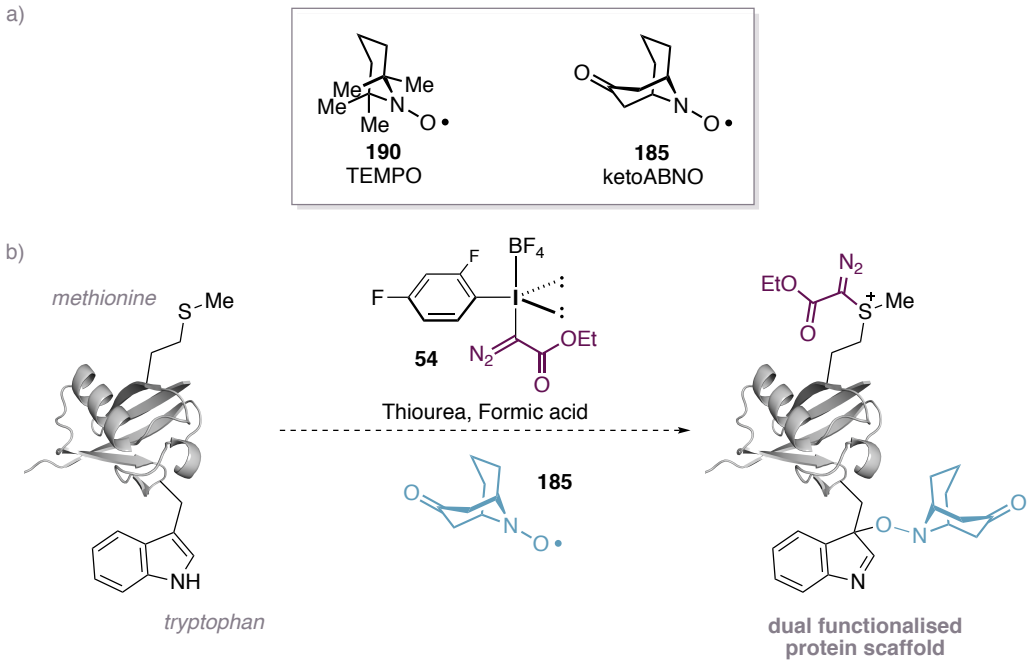
It was reported that the reactive species in the bioconjugation reaction was nitroxide **189**. **189** was thought to be formed *via* acid-promoted disproportionation of **185** and oxidation of resultant hydroxylamine **188** by NO_x, generated from $NaNO_2$ and acetic acid (Scheme 82). The authors proposed that subsequent bioconjugation proceeds *via* nucleophilic attack of the indole ring of a Trp residue towards the reactive nitroxide species **189**, followed by attack of water at the resultant imine.



Scheme 82 Proposed mode of reactivity of ketoABNO reagent with tryptophan residues

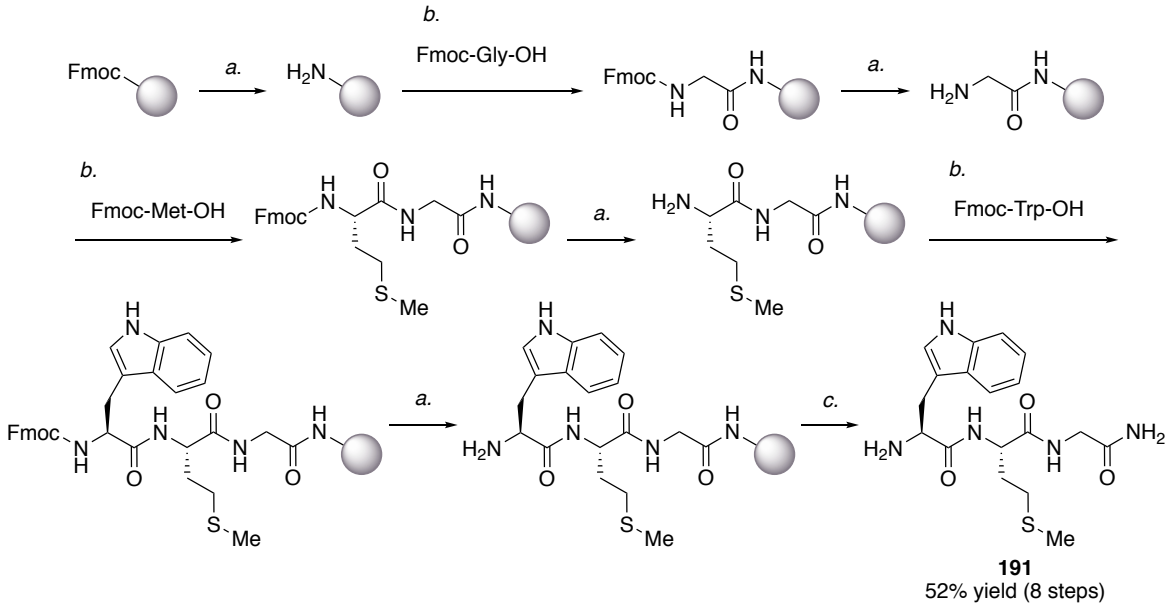
It was proposed that this tryptophan-selective bioconjugation strategy could be used in combination with methionine-selective bioconjugation for the synthesis of multi-functionalised protein scaffolds. An important additive for the methionine labelling procedure is TEMPO **190**, used as a scavenger of ROS to protect proteins against oxidative degradation during the reaction. It was noted that the organoradical **190**, and the tryptophan-selective organoradical **185**, were structurally similar (Scheme 83a). However, although structurally similar, Kanai had shown that TEMPO was unreactive in the Trp-labelling strategy. It was hypothesised that the additive **190**, used in the methionine strategy, could be replaced with the tryptophan-selective reagent **185**. Oxidation to the reactive nitroxide species **189** may occur *in situ* due to the fact that the reaction proceeds under oxidising conditions, with highly oxidising iodonium salt **54** present in a large excess. Subsequently, nucleophilic attack of the indole ring of a Trp residue to **189** could facilitate selective labelling of Trp, meanwhile iodonium salt **54** may selectively label Met using the established strategy.

If successful, this would provide a strategy for the simultaneous labelling of two different native amino acid residues, with two different functional groups, streamlining the synthesis of multi-functionalised protein scaffolds. Additionally, tryptophan and methionine are the two native amino acids with the lowest natural abundance in vertebrates (~1.3% and ~1.8% respectively).[23] As such, dual labelling of these residues may allow high selectivity in more complex systems.



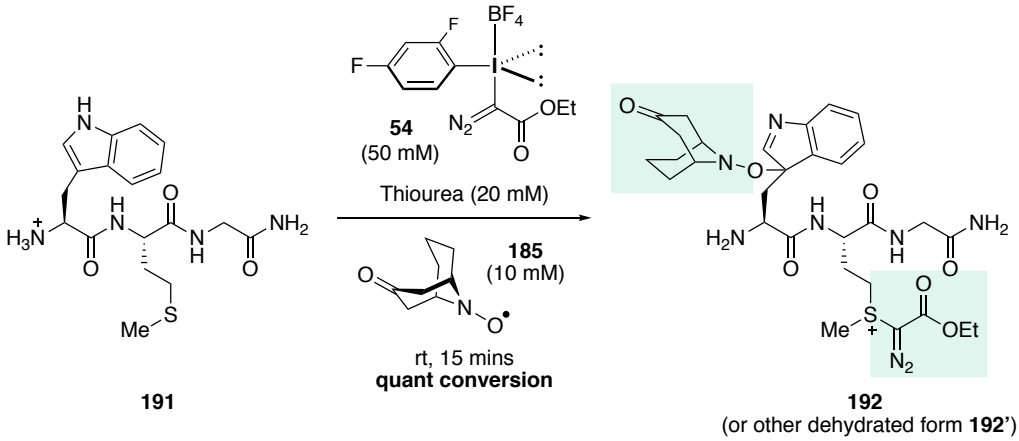
Scheme 83 a) Evident structural similarity between TEMPO and ketoABNO. b) Proposed simultaneous dual labelling of polypeptides and proteins using methionine labelling strategy, with substitution of TEMPO for Trp-selective reagent ketoABNO.

To initiate the development of the simultaneous dual labelling of methionine and tryptophan, a tripeptide model system was designed and synthesised using solid phase peptide synthesis so that the reaction could be studied on a small molecule system (Scheme 84). Sequential deprotection and amide coupling steps were carried out using Rink amide resin to facilitate the synthesis of tripeptide **191**. Final cleavage of the peptide from the resin afforded peptide **191** in an overall yield of 52% over 8 steps.



Scheme 84 Solid phase peptide synthesis of tripeptide **191**. *a*) 20% piperidine in DMF, rt, 15 mins. *b*) HCTU (1.5 eq.) DIPEA (1.5 eq.), rt, 1 hour. *c*) Solution of TFA (92.5%), triisopropylsilane (2.5%), CH_2Cl_2 (2.5%), H_2O (2.5%), rt, 1 hour.

Tripeptide was subjected to modified conditions for methionine bioconjugation, developed in Section 2.4, with the replacement of the additive TEMPO with ketoABNO, **185**, and a reaction time of 15 minutes (Scheme 85). Pleasingly, the dehydration product **192** (or other regioisomer **192'**) was observed in >95% conversion by LCMS analysis. This exciting result indicated that the dual labelling could be achieved in a single step from a native peptide, allowing the rapid formation of structurally complex peptidic scaffolds.



Scheme 85 Initial conditions for simultaneous dual functionalisation of Met and Trp on a model tripeptide substrate

Following this promising initial result, isolation of the mixture of dual labelled products was carried out in order to obtain MS/MS analysis. MS/MS analysis of the products of the reaction was required to confirm the two sites of modification in the tripeptide and ensure the selectivity of both reagents being used in tandem. As a number of labile bonds were present in the tripeptide product (e.g. N–O bond of ketoABNO motif), MS experiments were carried out at a low voltage of 10 V in the hope of observing fragment ions corresponding to the intact labelled amino acid residues. As discussed in Section 2.4.5, upon ionisation, alkylated methionine residues can fragment at the Cγ-SR_2 bond to produce an $[M-(SR_2-H)+H]^+$ ion in addition to the formation of the parent molecular ion $[M]^+$.[188,232,233] This fragmentation was evident in the MS/MS spectrum, and provided evidence that the diazoester species had transferred to the methionine residue as anticipated. Additional fragment ions corresponding to a number of bond cleavages were observed, depicted in Figure 16. Most notable were the y_1-(SMeR) and the -Trp fragments, corresponding to the loss of the modified Trp residue and modified Trp side chain respectively (Figure 16). This provided evidence that the ketoABNO reagent had reacted selectively with the indole of the tryptophan residue.

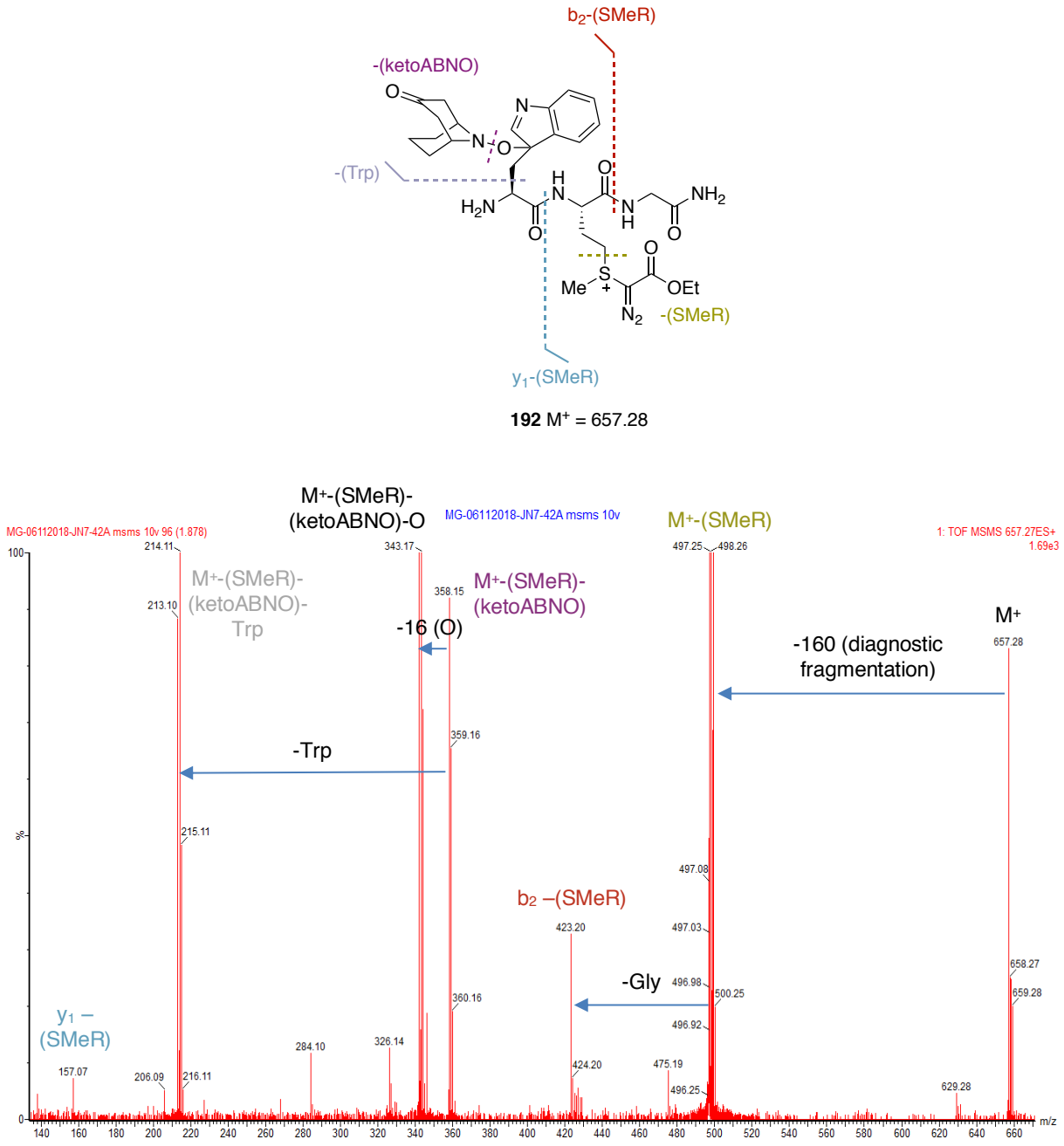
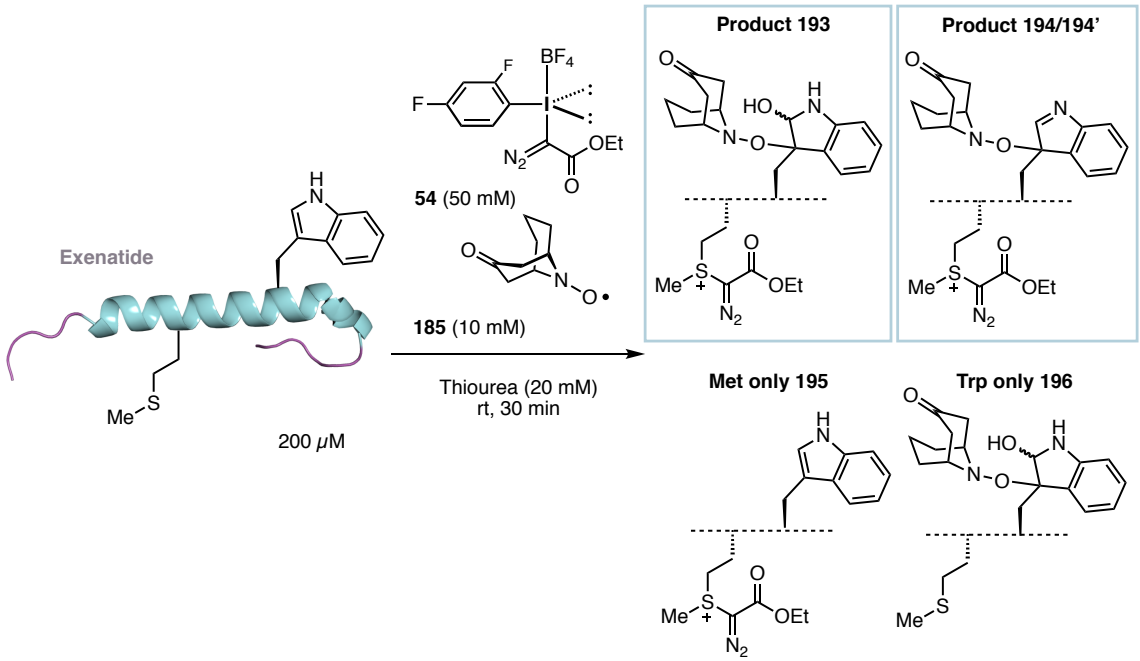


Figure 16 MSMS analysis of tripeptide **192** with diagnostic bond fragmentation denoted and colour coded

Subsequent efforts focused on the application of this novel strategy to larger polypeptide architectures. The polypeptide exenatide, of molecular weight 4186.6 Da, contains one methionine and one tryptophan residue and hence provided a suitable model polypeptide substrate for dual-labelling. Exenatide was subjected to the conditions used on the tripeptide system, using iodonium salt **54** (50 mM), ketoABNO **185** (10 mM) and thiourea (20 mM) at room temperature, with a reaction time of 30 minutes (Table 10).

The desired doubly-labelled products **193**, **194** and **194'** were observed in a combined conversion of 78%, as determined by LCMS total ion count. 4% of undesired product **195** was also observed, in which only the methionine residue had been labelled. Additionally, 18% of an oxidation side-product was observed, resulting from the dual-labelling and further oxidation of the peptide at an undetermined position. Nevertheless, high overall conversion to desired products was observed in 78% conversion on a complex scaffold. This preliminary result was encouraging and provided an excellent proof-of-concept for the use of this simultaneous labelling procedure on a polypeptide system.

Table 10 Initial attempt for the simultaneous dual labelling of Met and Trp residues in model polypeptide exenatide



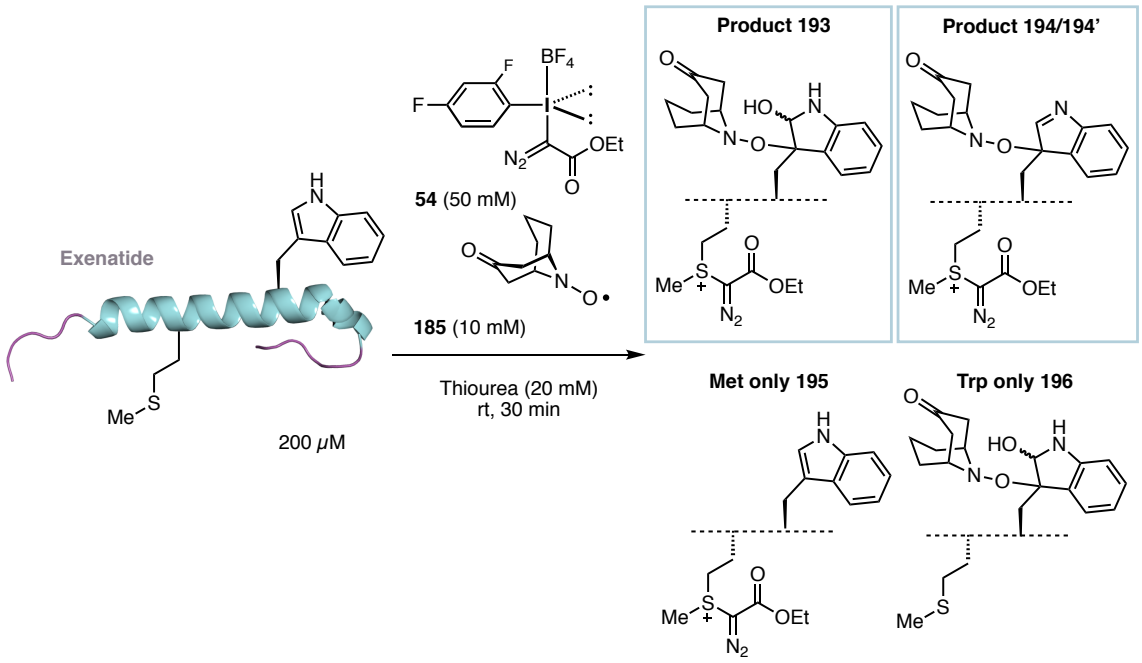
Entry	%conversion of exenatide	%product 193+194/194'	%oxidized peptide*	%Met only 195	%Trp only 196
1[a]	>95	78	18	4	0

[a]Reaction outcome determined by TIC. *Dual labelled product+[O]

Unexpectedly, on repeating the dual labelling of exenatide, under identical reaction conditions, varied results were obtained, with a combined conversion to desired products **193**, **194** and **194'** ranging from 16-55% (Entries 2-7, Table 11). Multiple attempts were made at reproducing the initial result of 78% conversion (Entry 1) without success. It seemed that irreproducibility was a problem with this dual functionalisation technique, with varied levels of oxidative degradation appearing to be the predominant inconsistency between identical reactions. Additionally,

incomplete tryptophan functionalisation also appeared problematic, with methionine labelled product **195** observed in dramatically different quantities (4-79%, Entries 1-7).

Table 11 Identical repeat reactions for the simultaneous dual labelling of Met and Trp residues in exenatide



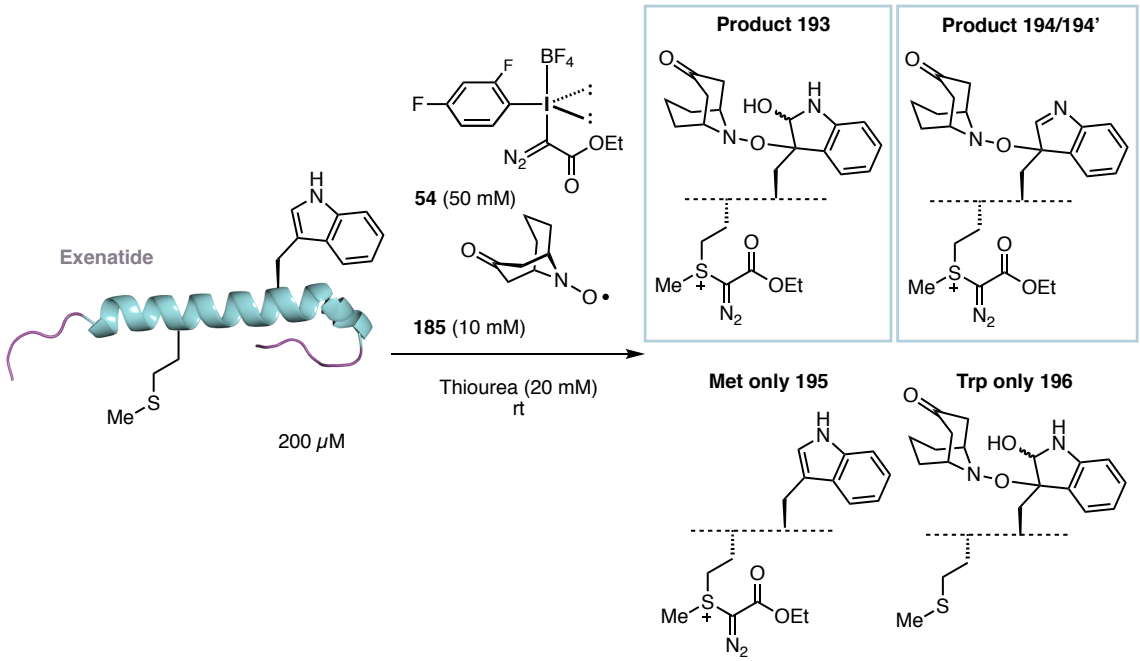
Entry	%conversion of exenatide	%product 193+194+194'	%oxidized peptide*	%Met only 195	%Trp only 196
1[a]	>95	78	18	4	0
2[a]	>95	55	4	41	0
3[a]	>95	48	2	49	0
4[a]	>95	51	4	44	0
5[a]	>95	45	4	50	0
6[a]	>95	53	14	32	0
7[a]	>95	16	5	79	0

[a]Reaction outcome determined by TIC. *Dual labelled product+[O]

Efforts to address the issue of unpredictable oxidative side-product formation began with carrying out the reaction under an inert atmosphere and using solvents degassed with N_2 (Table 12). As a large proportion of the oxidation observed was of the desired doubly-labelled species, it was envisaged that, under N_2, it may be possible to reduce oxidative degradation of dual-labelled products, improving on the conversion to **193**, **194** and **194'**. Unfortunately, an inert

atmosphere and N_2-degassed solvents provided no advantageous effect on conversion to products, with identical reaction conditions still resulting in inconsistent levels of oxidation (12-29%, Entries 1 & 2). On shortening the reaction time to 20 minutes, methionine-labelled product **195** was observed in only 25% conversion (Entry 3). Unexpectedly, increasing the reaction time to 60 minutes under a N_2 atmosphere led to reduced dual-labelling, with singly-labelled product **195** observed in 77% (Entry 4). Given that these attempts under a N_2 atmosphere had little positive effect on oxidation, for operational simplicity subsequent screening was carried out under air.

Table 12 Efforts to reduce oxidative side-products through N_2 atmosphere and N_2-degassed solvents at various reaction times



Entry	N_2/air	Reaction time	%conversion of exenatide	%product 193+194+194'	%oxidized peptide*	%Met only 195	%Trp only 196
1[a]	N_2	30 min	>95	88	12	0	0
2[a]	N_2	30 min	>95	66	29	5	0
3[a]	N_2	20 min	>95	65	10	25	0
4[a]	N_2	60 min	92	15	0	77	0

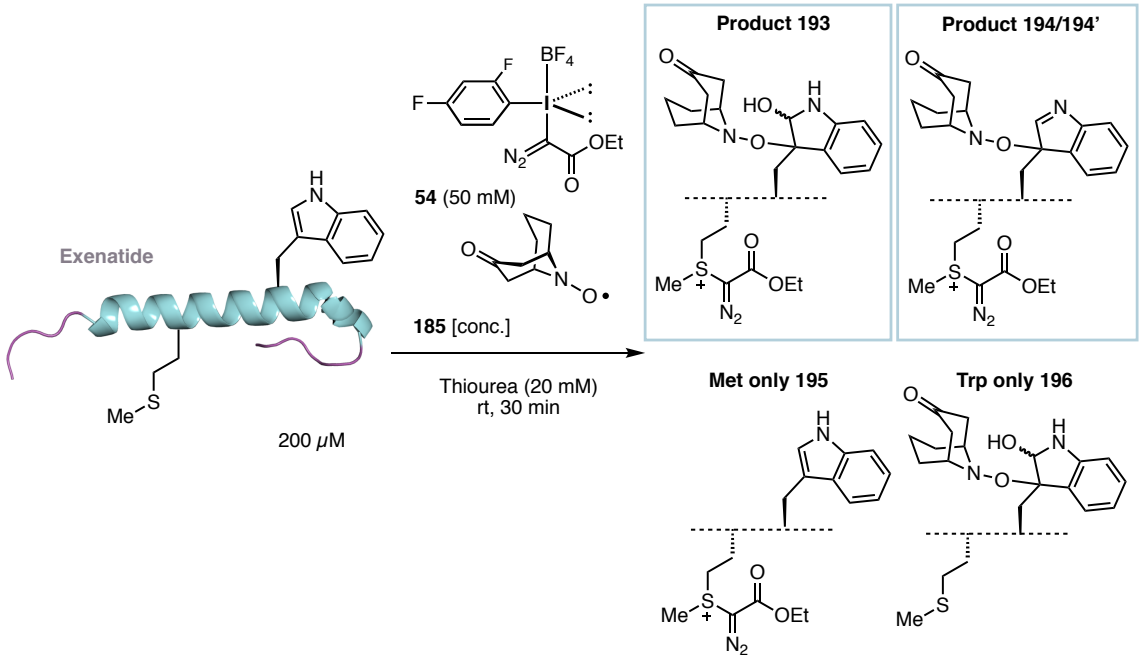
[a]Reaction outcome determined by TIC. *Dual labelled product+[O]. All reactions carried out using solvents degassed with N_2.

Efforts to improve the efficiency of the Trp labelling began by varying the amount of the Trp-selective ketoABNO reagent **185** in the reaction. Increasing the amount of ketoABNO used in

the reaction, from 10 mM to 40 mM, caused a decrease in conversion to desired dual-labelled products (Entries 1-3, Table 13). The decrease in yield corresponded to an increase in observed levels of undesired oxidation, with 42% conversion to oxidative side-products observed at a ketoABNO concentration of 40 mM (Entry 3). Decreasing the concentration of ketoABNO incrementally from 5 mM to 1 mM also had a detrimental effect on conversion to desired products **193**, **194** and **194'** (Entries 4-6). At a lower concentration of the Trp-selective reagent **185**, the labelling of Trp was poor, with <17% conversion to dual labelled products observed at concentrations <10 mM).

In the knowledge that TEMPO is unreactive towards tryptophan but can act as a scavenger of ROS, the amount of ketoABNO was reduced to 1 equivalent (with respect to the polypeptide) and TEMPO introduced as an additive (10 mM). Unfortunately, this achieved almost quantitative conversion to the methionine-labelled product **195**.

Table 13 Effect of varying the amount of ketoABNO used in dual functionalisation of Met and Trp on exenatide

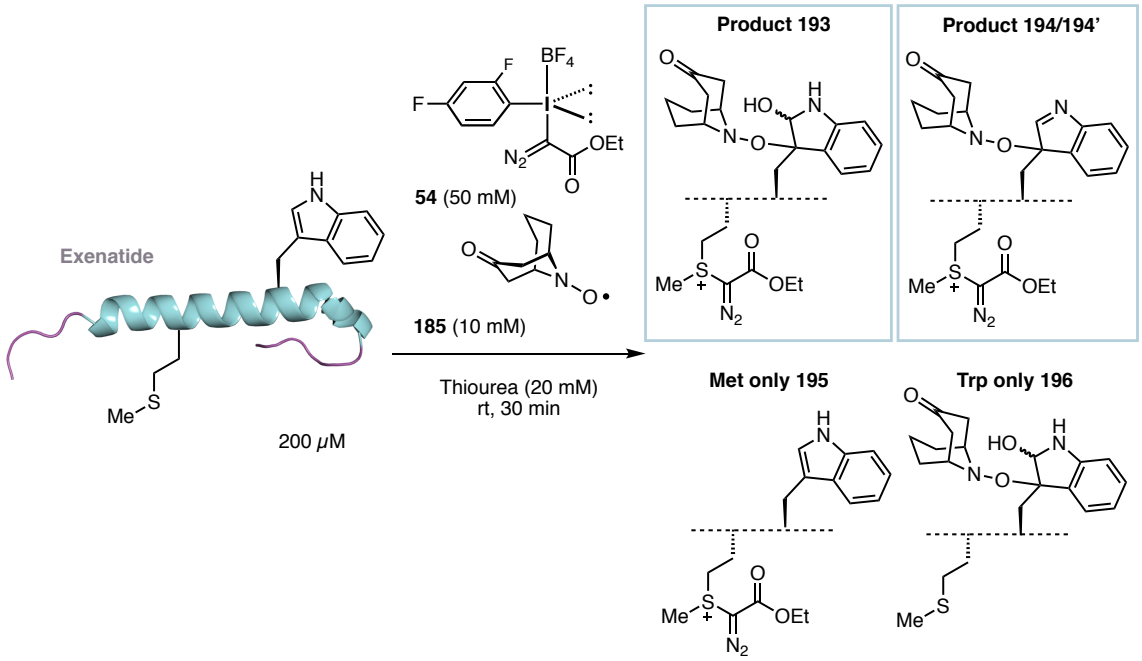


Entry	[ketoABNO]	%conversion of exenatide	%product 193+194+194'	%oxidized peptide*	%Met only 195	%Trp only 196	Comments
1[a]	20 mM	>95	52	12	36	0	-
2[a]	30 mM	>95	30	47	24	0	-
3[a]	40 mM	>95	32	42	27	0	-
4[a]	5 mM	>95	17	2	82	0	-
5[a]	2 mM	>95	4	0	96	0	-
6[a]	1 mM	>95	0	2	98	0	-
7[a]	0.2 mM	>95	0	0	>95	0	*10 mM TEMPO additive*

[a)]Reaction outcome determined by TIC. *Dual labelled product+[O]

Kanai and co-workers reported that the 'reactive species' in the tryptophan-selective bioconjugation reaction was the oxidised nitroxide species formed from the organoradical **185**. Therefore, it was proposed that pre-stirring of the ketoABNO organoradical with the oxidising iodonium salt species **54** may facilitate the formation of the reactive nitroxide species, prior to exposure to the polypeptide. Therefore, pre-stirring of **185** and **54** was carried out, under N_2, before the subsequent addition of this pre-stirred solution to the reaction mixture (Table 14). This pre-stirring technique gave inconsistent and irreproducible results and did not seem to correlate with improved conversion to products. For example, pre-stirring **185** and **54** for 10 minutes, under identical conditions, gave dramatically different results ranging from 0-88% conversion to desired products (Entries 5 & 6).

Table 14 Dual labelling of Met and Trp residues in exenatide with pre-stirring of iodonium salt and ketoABNO reagents, under N_2

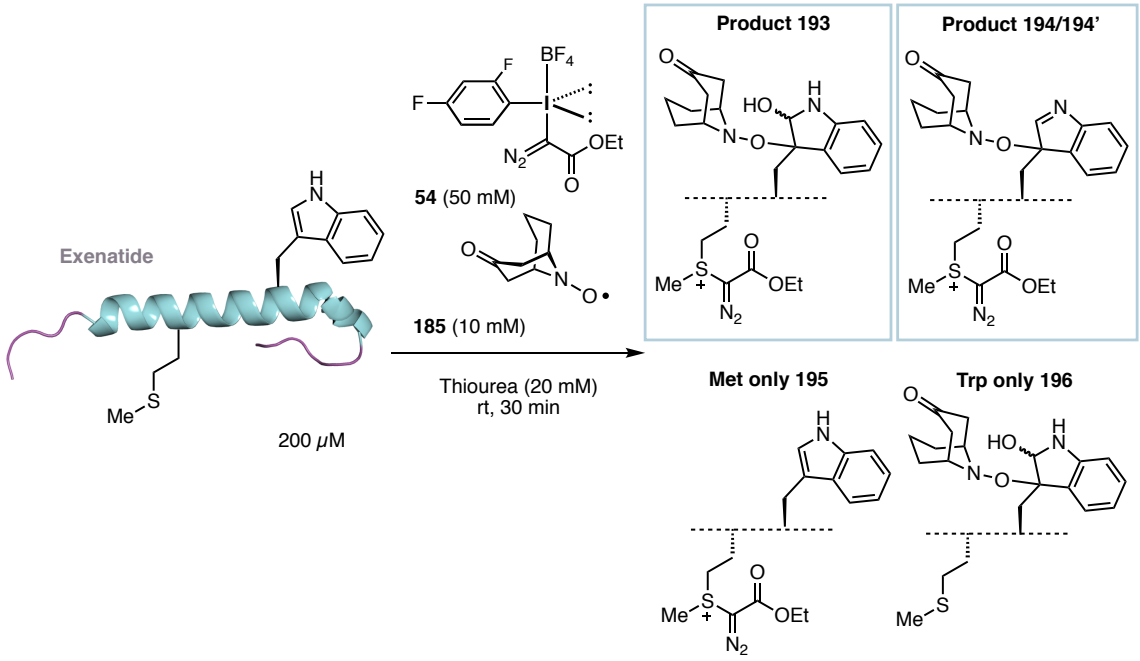


Entry	185 & 54 pre-stirring	%conversion of exenatide	%product 193+194+194'	%oxidized peptide*	%Met only 195	%Trp only 196
1[a,b,c]	1 min	>95	83	17	0	0
2[a,b,c]	1 min	>95	32	68	0	0
3[a,b,c]	5 min	95	43	0	52	0
4[a,b,c]	5 min	>95	31	69	0	0
5[a,b,c]	10 min	>95	88	0	12	0
6[a,b,c]	10 min	>95	0	0	>95	0

[a]Reaction outcome determined by TIC. [b]Reaction time 30 minutes. [c]Reaction conducted under N_2, using N_2 degassed solvents *Dual labelled product+[O]

Pre-stirring in the same manner was conducted under air and obtained similarly inconsistent results (Table 15). Varying levels of conversion to desired products was observed, with no apparent correlation between pre-stirring reagents and improvement of the dual-labelling or decrease in formation of undesired oxidation side-products.

Table 15 Dual labelling of Met and Trp residues in exenatide with pre-stirring of iodonium salt and ketoABNO reagents, under air



Entry	185 & 54 pre-stirring	%conversion of exenatide	%product 193+194+194'	%oxidized peptide*	%Met only 195	%Trp only 196
1[a,b]	1 min	>95	64	36	0	0
2[a,b]	5 min	>95	47	0	53	0
3[a,b]	10 min	83	41	8	42	0

[a]Reaction outcome determined by TIC. [b]Reaction time 30 minutes. *Dual labelled product+[O]

In order to observe the formation of products over the course of the reaction, a larger scale reaction was carried out, under air, in which aliquots were acquired and analysed periodically (Figure 17). The rate of dual-labelled product formation was rapid, with quantitative conversion of starting material observed after only 1 minute, forming ~80% of desired dual-labelled products **193**, **194** and **194'** and ~20% oxidation. At this point it was clear that the reaction was complete, with product ratios remaining constant for the duration of the analysis. No singly-labelled product was observed, indicating that the tryptophan labelling could also occur at a similar rate to that of the methionine labelling.

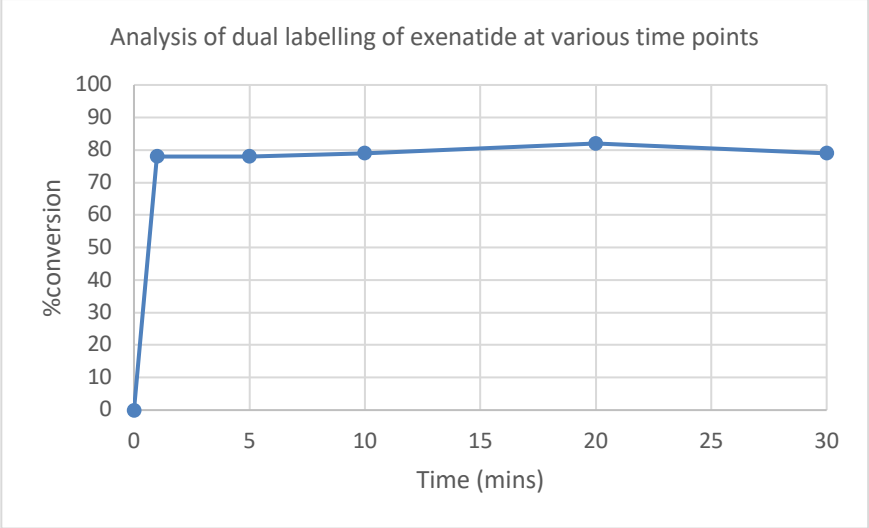
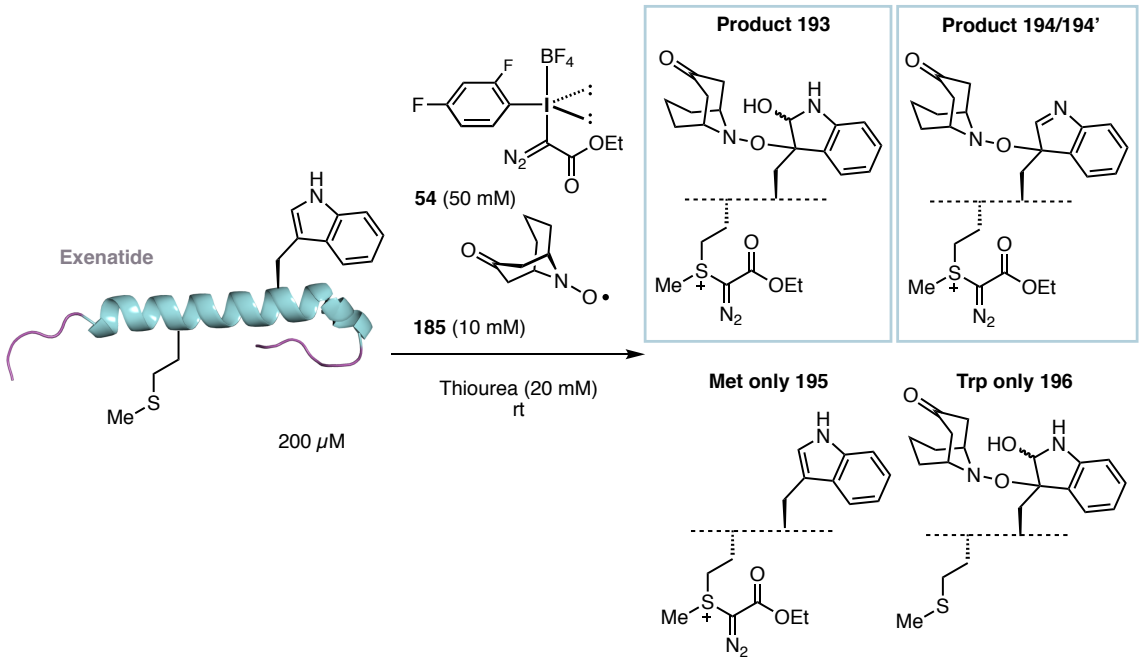


Figure 17 Monitoring of the conversion to products for the dual labelling of exenatide. % conversion determined of aliquots acquired at time points throughout the reaction using LCMS TIC.

It had previously been observed that oxidation had occurred over reaction times greater than 1 minute, the time in which product formation had been observed. Therefore, it was proposed that rapid formation of the desired products could be impeded by the subsequent oxidative degradation of products occurring in the highly oxidising environment of the reaction over time. As a result, the reaction time was shortened and analysed after 1 minute (Table 16). Under identical conditions the reaction remained capricious, providing variable conversions to the desired products from 50-81% (Entries 1-4). Undesirable oxidation of the dual-labelled product remained problematic, with the oxidative side-product formed in 19-50% conversion.

Table 16 Screening of dual labelling of Met and Trp on exenatide with a shortened reaction time



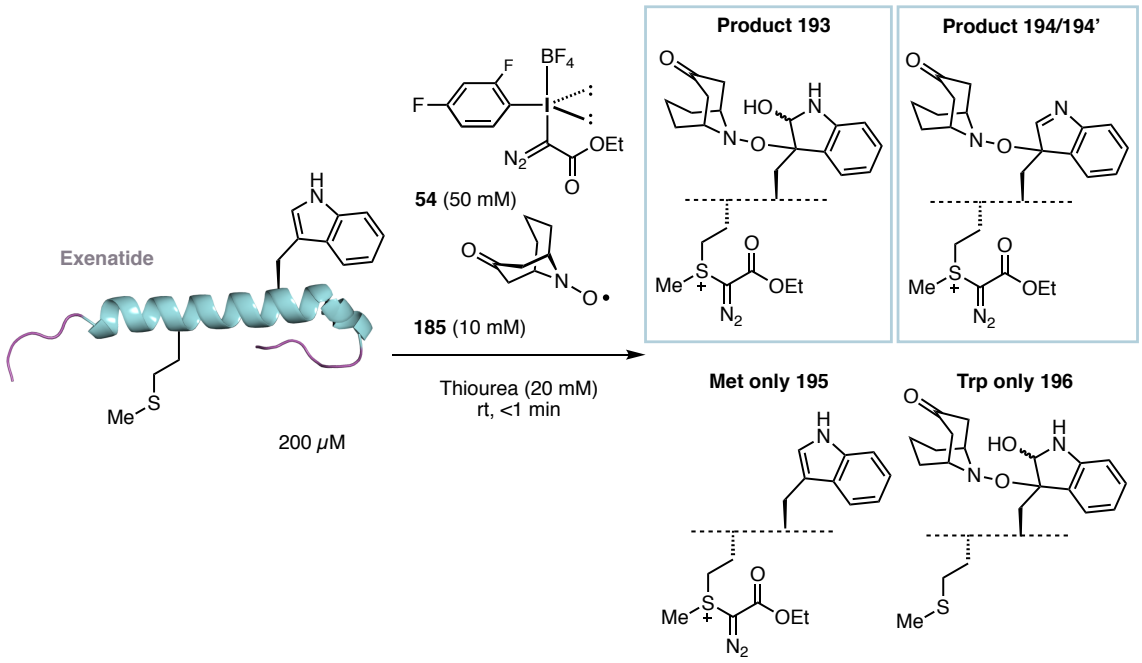
Entry	Reaction time	%conversion of exenatide	%product 193+194+194'	%oxidized peptide*	%Met only 195	%Trp only 196
1[a]	1 min	>95	81	19	0	0
2[a]	1 min	>95	68	32	0	0
3[a]	1 min	>95	50	50	0	0
4[a]	1 min	>95	75	25	0	0

[a)] Reaction outcome determined by TIC. *Dual labelled product+[O]

It was hypothesised that quenching the reaction with an antioxidant may reduce the amount of oxidation side-product formed. In light of the rapid formation of the dual labelled products, it was proposed that a short reaction time followed by a reductive quench may circumvent the oxidative degradation and afford high conversion to the desired products. The dual-labelling procedure was conducted and the reaction was quenched using ascorbic acid, a common antioxidant which is known to interact directly with oxidising radical species[164] (Table 17). Interestingly, the use of ascorbic acid completely halted the formation of oxidation side-products (Entries 1-4). Depending on the reaction time prior to quenching, the desired dual labelled products **193**, **194** and **194'** were formed in a combined conversion of 43-81%, with the undesired singly-labelled product observed in 8-45% conversion. These promising results for the use of ascorbic acid indicated that the levels of oxidation could be controlled. However,

considering the tryptophan labelling was not reaching completion, the conversion to dual labelled products now required improvement.

Table 17 Investigation into the effect of quenching the dual labelling reaction with an antioxidant/basic conditions



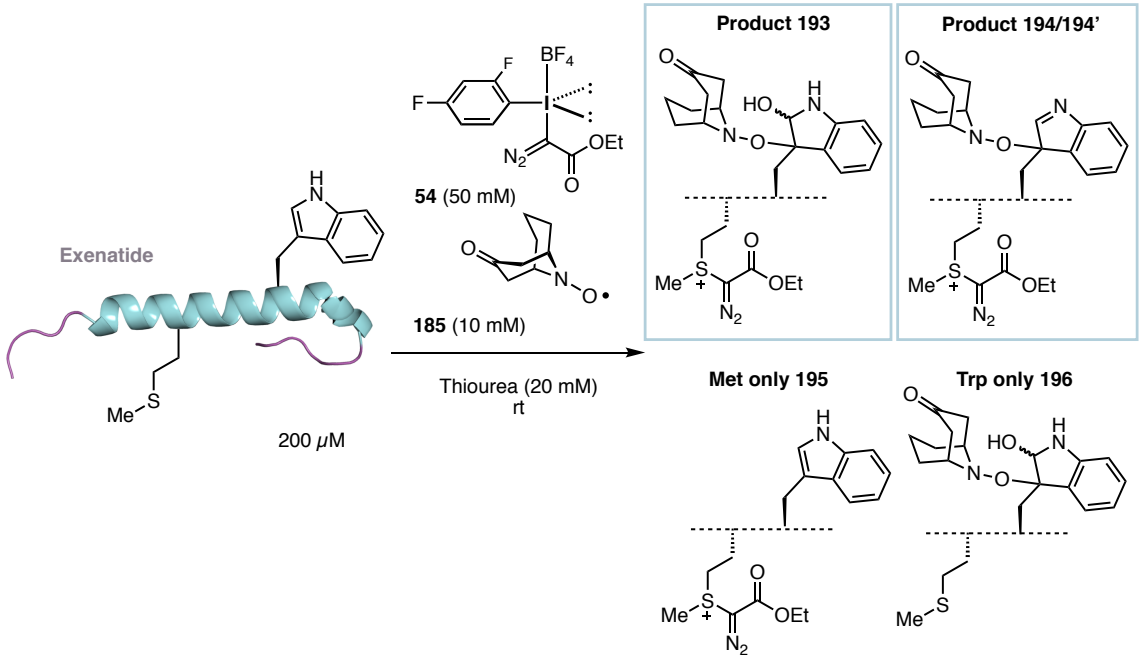
Entry	Quenching conditions	Reaction time	%conversion of exenatide	%product 193+194+194'	%oxidized peptide*	%Met only 195	%Trp only 196
1[a]	ascorbic acid (60 mM)	Immediate quench	88	43	0	45	0
2[a]	ascorbic acid (60 mM)	15 sec	>95	75	0	18	0
3[a]	ascorbic acid (60 mM)	30 sec	>95	81	0	14	0
4[a]	ascorbic acid (60 mM)	60 sec	>95	77	0	8	0

[a] Reaction outcome determined by TIC. *Dual labelled product+[O]

Having circumvented the problematic oxidation which was occurring during longer reaction times, it was proposed that improved conversion could be achieved by lengthening the reaction time, potentially allowing full labelling of the tryptophan residue. Reactions were screened with durations ranging from 30 seconds to 30 minutes before quenching with ascorbic acid (60 mM) (Table 18). Pleasingly, no undesired oxidation side-products were observed and increasing the reaction time improved the conversion to dual labelled products slightly, affording the desired dual-labelled products in 79% conversion over 10 minutes (Entry 5). Further increase of the reaction time provided negligible improvement and after 30 minutes products resulting from oxidation were observed (Entries 6 & 7). Considering these results, it was determined that a

reaction time of 10 minutes provided optimal conversion to dual-labelled products. Singly-labelled side-product **196** was also observed at 7-9% conversion under these reaction conditions, in which the methionine labelling had not gone to completion.

Table 18 Screening of reaction times for the dual labelling of Met and Trp residues in exenatide, quenching the reaction with ascorbic acid

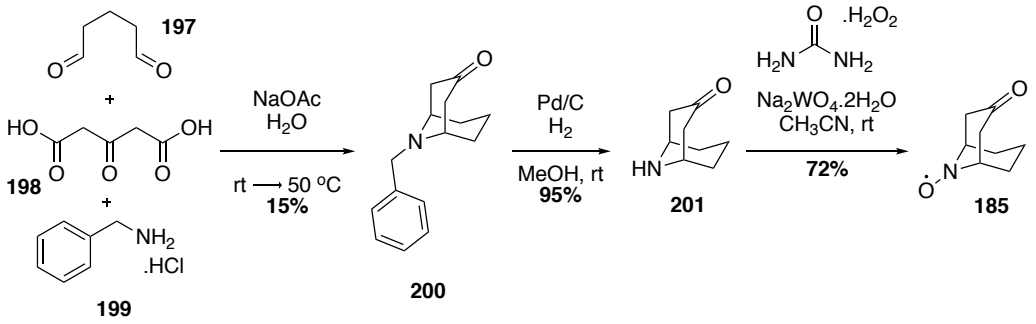


Entry	Reaction time	%conversion of exenatide	%product 193+194+194'	%oxidized peptide*	%Met only 195	%Trp only 196
1[a,b]	30 sec	>95	71	0	12	8
2[a,b]	1 min	>95	74	0	11	7
3[a,b]	2 min	>95	72	0	11	8
4[a,b]	5 min	>95	75	0	7	7
5[a,b]	10 min	>95	79	0	5	7
6[a,b]	20 min	>95	79	0	8	7
7[a,b]	30 min	>95	65	11	9	9

[a)] Reaction outcome determined by TIC. [b)]Quenched with ascorbic acid (60 mM) *Dual labelled product+[O]

The purity and origin of the ketoABNO reagent was also found to be important for the reproducibility of the dual labelling. Commercial sources possessed variable purity and stability. Therefore, the ketoABNO reagent was synthesised and recrystallised (Scheme 86).

Using freshly recrystallised material improved the reproducibility of results for the dual labelling. This pure material was used in the remainder of the optimisation.

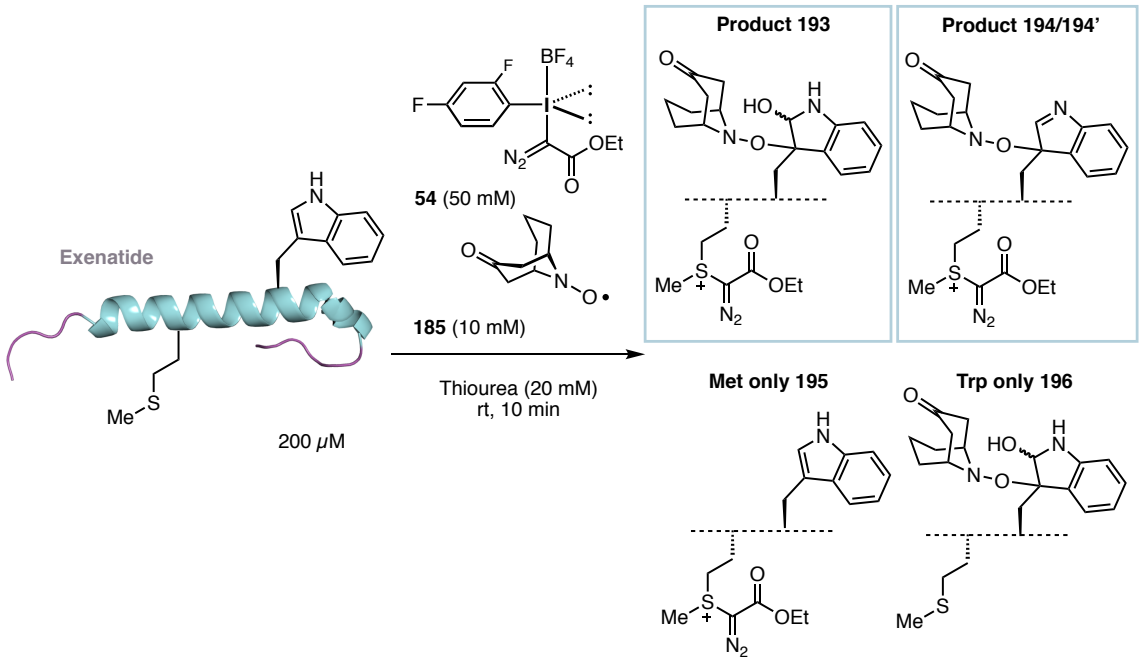


Scheme 86 Synthesis of ketoABNO **185**

Different concentrations of ascorbic acid were also screened to observe the effect on conversion to dual-labelled products (Table 19). It was found that on decreasing the amount of ascorbic acid used to quench the reaction, the conversion to dual labelled products **193**, **194** and **194'** improved, reaching 90% at a concentration of 1 mM ascorbic acid (Entry 6). In this case, no singly-labelled products were observed, with the labelling of both the methionine and the tryptophan residue having reached completion. Notably, at the lowest concentration of ascorbic acid screened, some non-specific labelling of the polypeptide by iodonium salt **54** was observed, with the incorporation of two equivalents of the diazoester motif to exenatide. The conditions yielding the highest conversion (Entry 6) were repeated and gave consistent conversions to desired products of 83-92% (n = 5), providing a reproducible strategy for dual functionalisation.

In the initial tryptophan labelling, reported by Kanai, tryptophan labelling was achieved in the range of 15-64% conversion for polypeptide and protein systems with greater than 4 amino acids. Using optimal conditions developed herein, the labelling of two different amino acids could be achieved on a 39 amino acid polypeptide system in 90% conversion. This demonstrates a vast improvement on the reported tryptophan labelling reaction, in addition to conducting the reaction simultaneously with the labelling of the methionine amino acid residue.

Table 19 Screening of concentration of ascorbic acid used for quenching of the dual labelling reaction



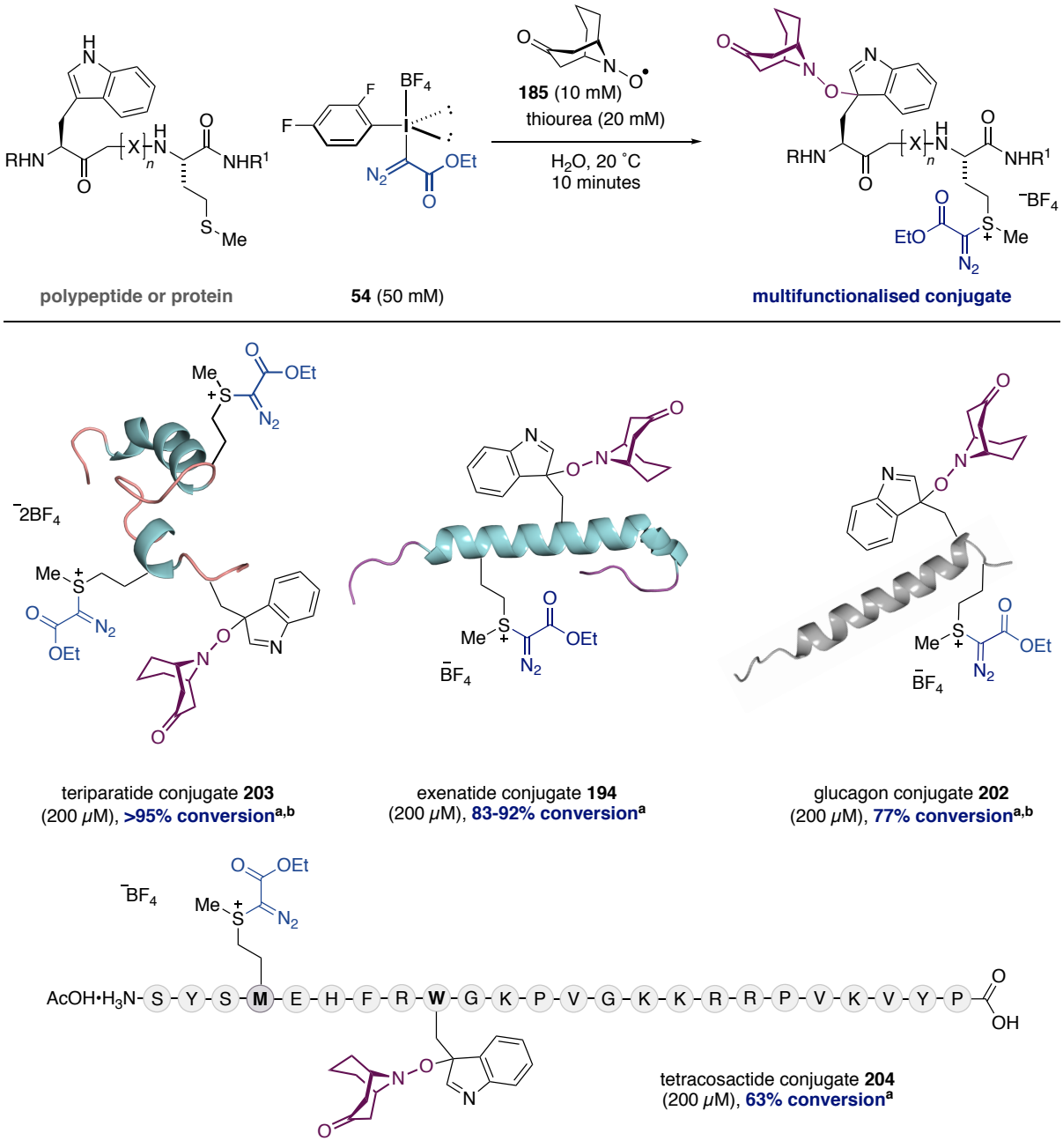
Entry	Ascorbic acid quench	%conversion of exenatide	%product 193+194+194'	%oxidized peptide*	%Met only 195	%Trp only 196
1[a,b,c]	60 mM	>95	72	0	11	9
2[a,b,c]	40 mM	>95	74	0	12	8
3[a,b,c]	20 mM	>95	78	0	11	6
4[a,b,c]	10 mM	>95	80	0	6	3
5[a,b,c]	5 mM	>95	80	0	9	3
6[a,b,c]	1 mM	>95	90[d]	0	0	0

[a)] Reaction outcome determined by TIC. [b)]Reaction time = 10 mins. [c)]Small amount of oxidised exenatide starting material observed. [d)]+9% product+113 observed (non-specfic labelling). *Dual labelled product+[O]

With optimised conditions in hand, attention turned to investigating the scope of polypeptide or protein substrates which could be efficiently labelled using this simultaneous functionalisation technique (Scheme 87). Glucagon is a 29 amino acid, helical polypeptide of molecular weight 3482.8 Da which contains one methionine and one tryptophan residue. It was found that this substrate could be labelled at both Met and Trp simultaneously, in 77% conversion to products **202**.

Teriparatide is the 34-amino acid, bioactive, N-terminal domain of parathyroid hormone and is used in the treatment of some forms of osteoporosis.[226] The sequence contains two methionine residues and one tryptophan residue, and the secondary structure includes two alpha helices.

Pleasingly, simultaneous Met and Trp labelling of teriparatide occurred in quantitative conversion to triply-labelled products **203**. Tetracosactide, a 24-amino acid synthetic peptide approved for medical use,[225] was also tolerated in the reaction. However, this substrate provided lower conversion to desired products **204**, with the reaction profile proving more complex.



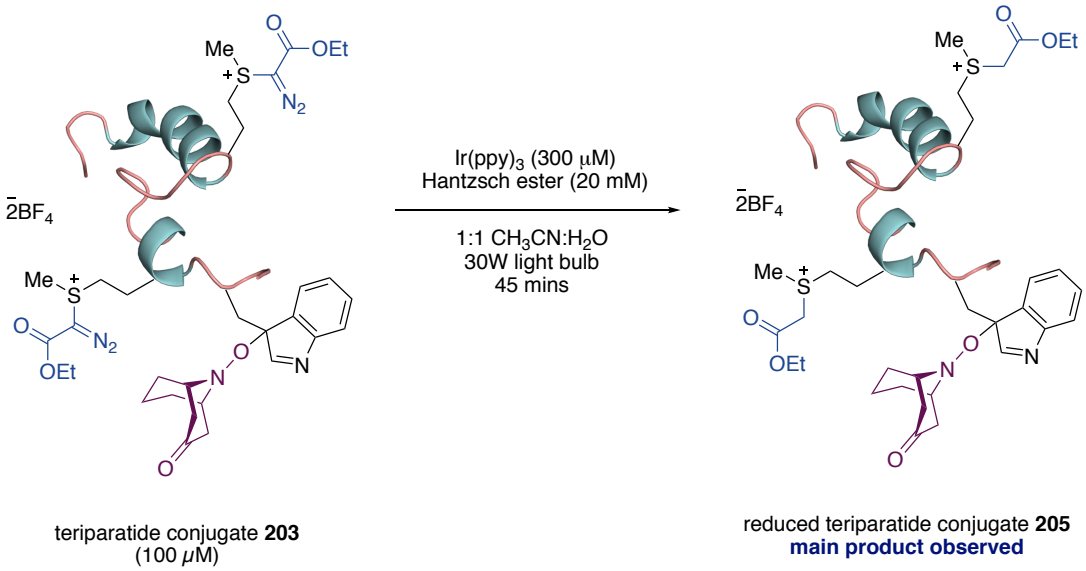
Scheme 87 Substrate scope for the multi-functionalisation of Met and Trp residues in polypeptides. Products are depicted as dehydration products but conversions include formation of hydration product + two dehydration products. [a]Conversion determined using LCMS TIC. [b]Formic acid (0.1 M) used in the reaction.

Isolation and further characterisation of these multi-functionalised polypeptide scaffolds was attempted. However, due to the fact that three inseparable products were formed during the

reactions (dual labelling product + two dehydration products), isolation proved difficult. As a result, MS/MS and CD analysis could not be carried out. However, MS/MS had been performed on a smaller peptidic system, confirming the selectivity of the dual labelling.

It had been established that reduction of the diazo moiety in the β-sulfonium α-diazoester improved the stability of the resulting sulfonium conjugates (Section 2.4.6.2). Kanai and co-workers reported that, on a small molecule system, the tryptophan modification was indefinitely stable under a variety of conditions and at various pH levels.[139] Therefore, it was proposed that subjecting the multi-functionalised polypeptide species to the photoreduction conditions may provide a facile strategy for improving conjugate stability.

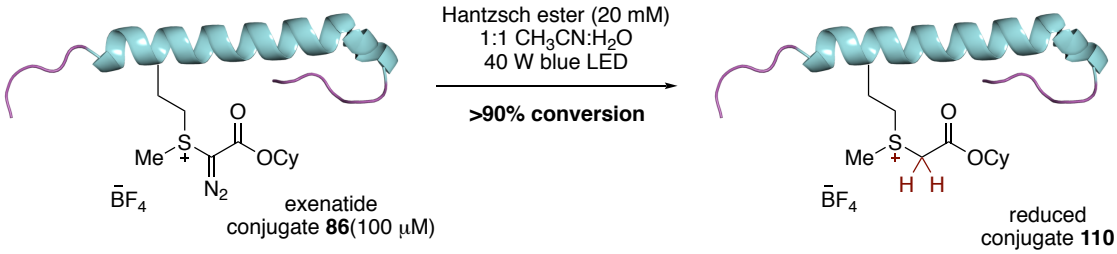
Exposing the triply-labelled teriparatide conjugate **203** to the photoreduction conditions, using $Ir(ppy)_3$ and Hantzsch ester, gave a promising preliminary result with the conjugate **205** observed as the major product, with both diazo motifs reduced (Scheme 88). This result, indicating that **203** is amenable to photoredox-mediated reduction of the diazo moiety, suggests that the stability of these highly modified scaffolds could be improved in an analogous fashion to the species singly-labelled at methionine (Section 2.4.6.2). Detailed stability studies on multi-functionalised polypeptide scaffolds are currently underway.



Scheme 88 Initial attempt at photoredox-mediated reduction of sulfonium diazo motif on triply modified teriparatide conjugate **203**

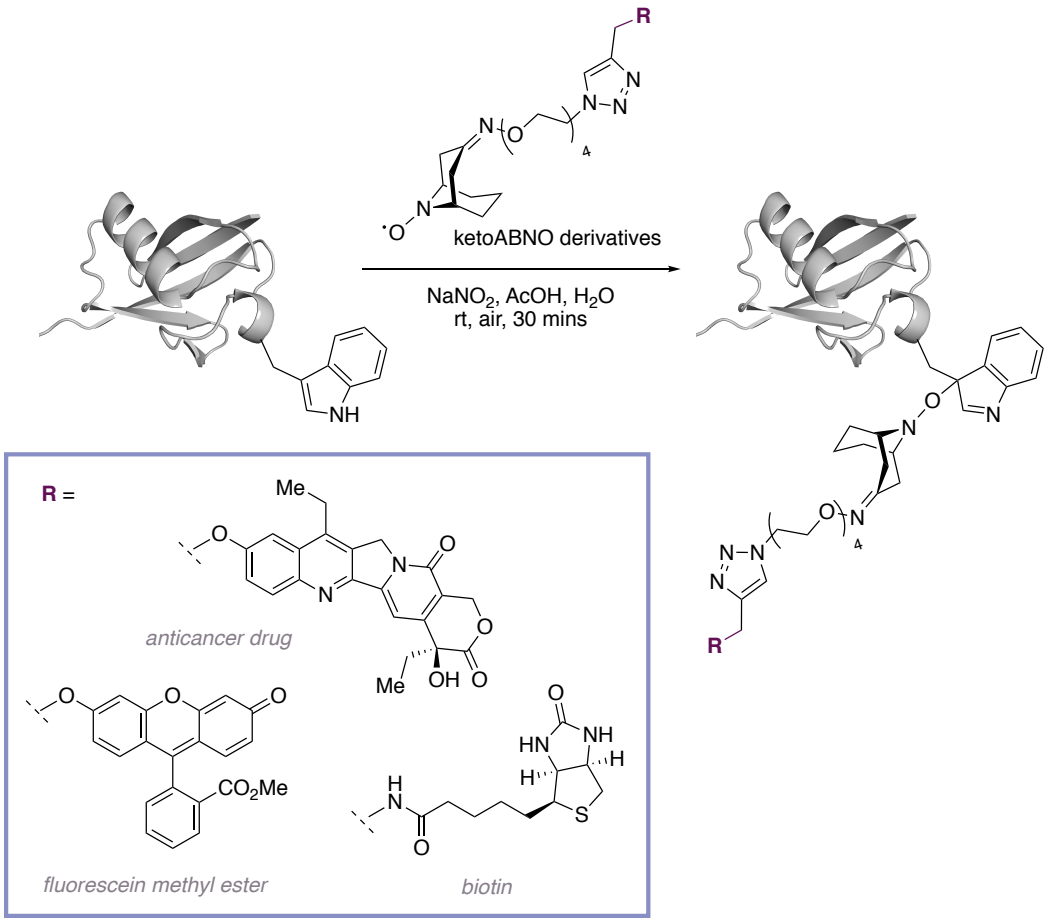
3.4 Future work

Moving forward with the simultaneous multi-functionalisation would require investigation into the stability of the β-sulfonium α-diazoester dual labelled conjugates to ensure a synthetically useful half-life which would enable their use in biologically relevant applications. The visible light-mediated reduction of these multi-functionalised scaffolds presented a promising preliminary result which may enable the stability of the β-sulfonium α-diazoester conjugates to be improved. Since the development of the initial visible light-mediated reduction, it has been discovered that the reaction can be conducted in the absence of the photocatalyst when using a more powerful, 40 W blue LED Kessil lamp (Scheme 89). Therefore, it would be interesting to explore these photocatalyst-free conditions on the multi-functionalised scaffolds as a means of improving their half-life.



Scheme 89 Photocatalyst-free, visible light-mediated reduction of exenatide β-sulfonium α-diazoester conjugate

Kanai and co-workers detailed various functional payloads which could be transferred to polypeptide and protein systems using their tryptophan-selective bioconjugation approach, including fluorescein, biotin and an anticancer drug molecule (Scheme 90).[139] It has also been established that various functional handles and payloads can be transferred to polypeptides and proteins using the methionine bioconjugation strategy, with functionalised iodonium salt reagents (Section 2.4.3). An important future development of the simultaneous multi-functionalisation would be investigation into the compatibility of these functionalised reagents with the dual-labelling strategy, exploring the different biologically relevant payloads which may simultaneously be transferred.



Scheme 90 Scope of functional payloads which could be transferred using the tryptophan-selective bioconjugation procedure developed by Kanai *et al.*[139]

Finally, it would be desirable to demonstrate the simultaneous multi-functionalisation on a larger protein system. Following identification of a suitable protein substrate in which, ideally, a single surface exposed methionine and tryptophan residue are present, it would be useful to demonstrate that the labelling is applicable to more complex protein scaffolds, bearing additional secondary structural features.

3.5 Summary

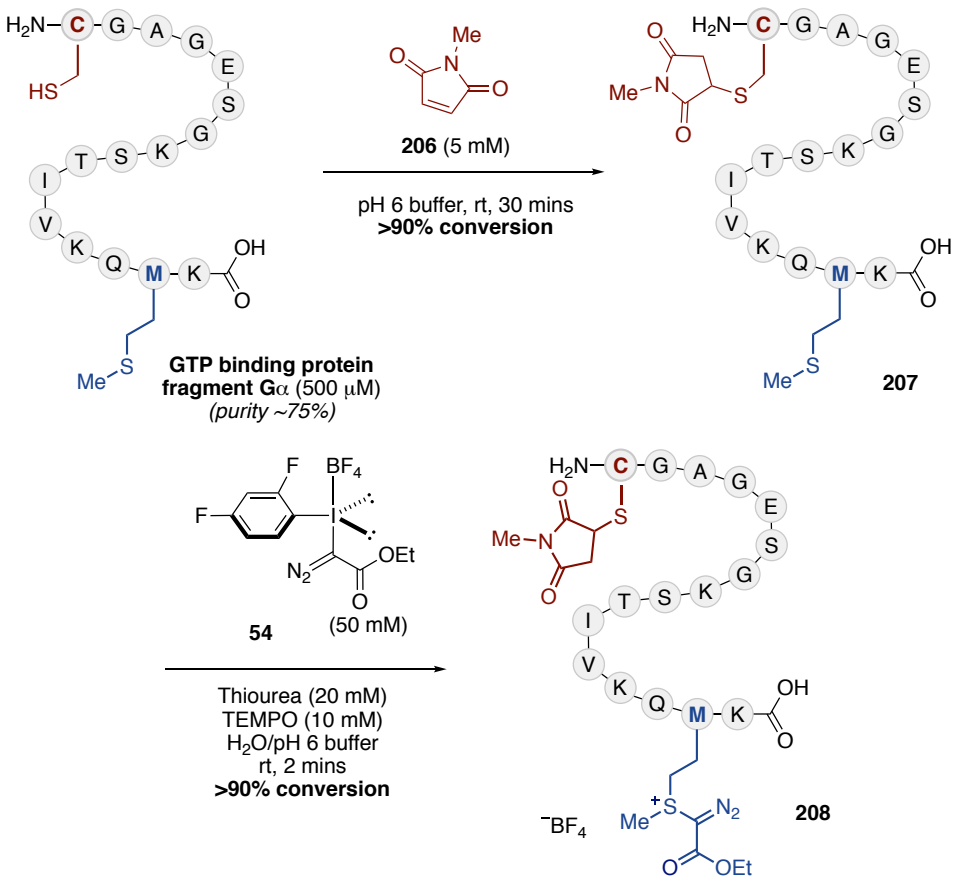
This chapter has described the application of the methionine bioconjugation strategy, developed in Section 2, to the synthesis of multi-functionalised polypeptide scaffolds. A reported tryptophan-selective bioconjugation technique was identified, which employed an organoradical species, ketoABNO, for indole functionalisation. This technique was utilised for the development of a selective dual-functionalisation strategy in combination with the methionine bioconjugation developed in Section 2. Using the optimised methionine labelling conditions, organoradical TEMPO was replaced with the structurally similar organoradical ketoABNO to create a simultaneous dual-functionalisation strategy. Extensive optimisation identified ascorbic acid as a key reagent in the simultaneous dual-labelling, which prevented the formation of undesired side-products resulting from oxidation of the polypeptide. Various polypeptides were tolerated in the reaction, facilitating the formation of dual- or even triply-modified polypeptide scaffolds in a single step from native polypeptides. Finally, a promising preliminary result identified that the photoredox-mediated reduction of the β-sulfonium α-diazoester (developed in Section 2.4.6.2) was applicable to multi-functionalised species, providing a potential strategy for improving the stability of resulting conjugates.

4 Conclusions and Outlook

Bioconjugation, the selective functionalisation of native chemical functionality within biomolecules, has emerged as an effective platform for the introduction of exogenous functional groups into polypeptides and proteins. The resulting functionalised scaffolds are exploited for myriad biological applications, including targeted drug delivery and cellular imaging. Many bioconjugation techniques have exploited the reactivity of the amino acids cysteine and lysine, whose nucleophilic side chains are susceptible to nucleophilic attack.[2] In contrast, strategies targeting the amino acid methionine for selective bioconjugation have remained underexplored.

To address this, a general strategy for methionine bioconjugation has been developed. Through the use of hypervalent iodine reagents, alkylation at the thioether side chain of methionine could be achieved selectively in under 5 minutes, forming a β-sulfonium α-diazoester moiety. The reaction was found to have exceptional functional group tolerance, although free cysteine residues were concomitantly oxidised to the corresponding disulfide.

While incompatibility with free cysteine residues is known for other oxidative functionalisation strategies,[192] it is acknowledged that overcoming this limitation would represent an important advance for methionine bioconjugation. Future work will involve the development of a more general platform for methionine functionalisation which can tolerate free cysteine residues. It has been established that through prior reaction of a free cysteine with *N*-methylmaleimide **206**, a methionine residue in a 16 amino acid synthetic peptide can be selectively functionalised using iodonium salt **54** without cysteine oxidation (Scheme 91). Interestingly, iodonium salt **54** displayed selectivity for methionine over the thioether present at the functionalised cysteine residue of **207**. This not only allows the selective formation of a desirable dual-labelled scaffold (see Section 3), it presents a strategy for circumventing the incompatibility of the methionine bioconjugation with free cysteine residues. It would be desirable to find a more general solution to this problem and efforts to this end may involve exploring, for example, a milder hypervalent iodine reagent which could result in less deleterious oxidation.

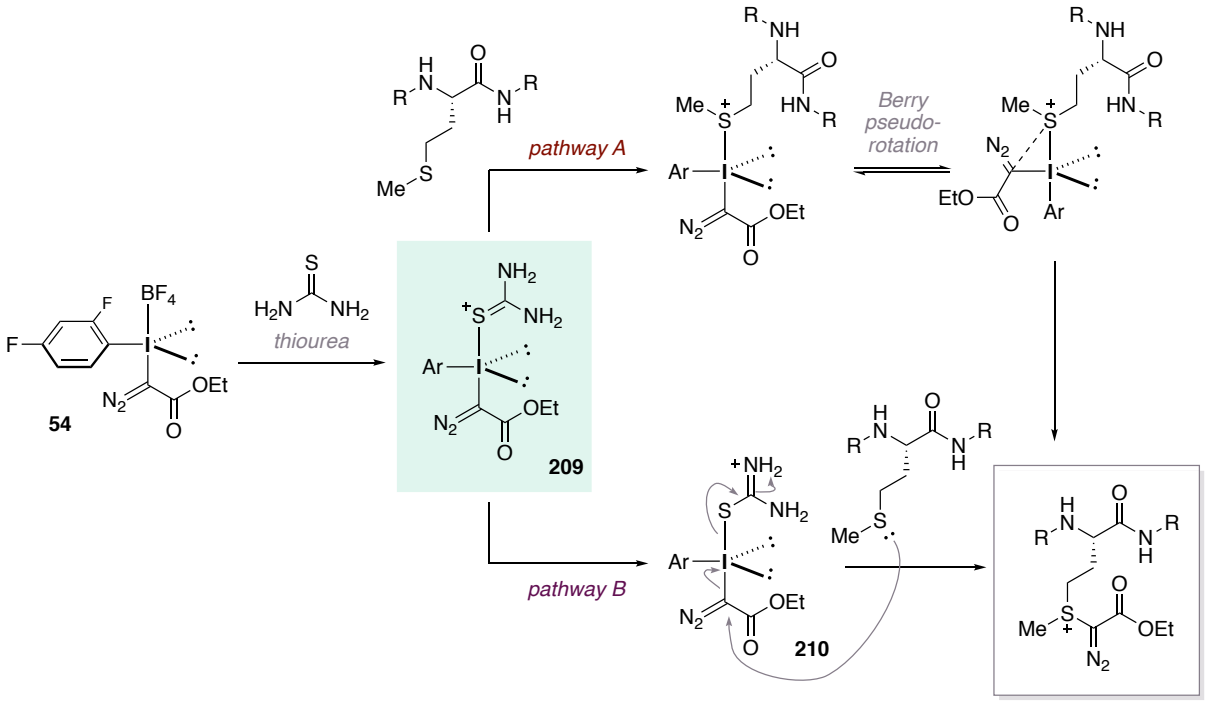


Scheme 91 Sequential dual labelling of GTP binding protein fragment Gα exploiting cysteine-maleimide chemistry and methionine bioconjugation. Conversions determined by LCMS TIC. Starting material was acquired in ~75% purity, therefore conversions were determined excluding the residual impurity carried through the synthesis.

Various functional groups and biologically relevant payloads were shown to be transferred effectively to the biomolecule of interest using this bioconjugation method, through the synthesis and exploitation of functionalised hypervalent iodine reagents. It was observed that alteration of the ester group of the iodonium salt reagent had an effect on the stability of the sulfonium products, modulating the rate of ester hydrolysis. Future endeavours aiming to exploit ester hydrolysis under stability control may be advantageous as a controlled drug delivery mechanism, with a pharmaceutical agent appended to the ester payload.

The selective methionine labelling could be applied to numerous polypeptide and protein substrates, however was attempted only on commercially available proteins bearing a surface exposed methionine residue. It would be desirable to demonstrate that larger protein scaffolds were tolerated in the reaction, incorporating more complex tertiary or even quaternary structural elements. Additionally, it would be interesting to determine the compatibility of this bioconjugation technique for the labelling of a monoclonal antibody, engineered to display surface exposed methionine residues, for the synthesis of antibody-drug conjugates.

The mechanism for the functionalisation of methionine using hypervalent iodine reagents is currently not well understood. It has been observed that thiourea plays a key role in the functionalisation, with a rate enhancement effect enabling labelling of methionine residues in <5 minutes. Speculative proposals for the role of thiourea include its role in the formation of a reactive intermediate species **209**, which could potentially be displaced by methionine followed by reductive ligand coupling (pathway A, Scheme 92), or direct nucleophilic attack of methionine at the diazo carbon of intermediate **210** (pathway B, Scheme 92). Investigations to elucidate the mechanism of the reaction will form an important part of ongoing research. It is envisaged that experimental and computational studies will assist in the mechanistic investigation and may give an insight into likely intermediates involved in the reaction.

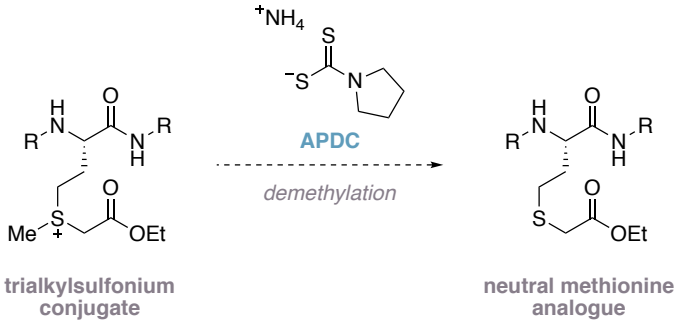


Scheme 92 Potential mechanistic pathways for the functionalisation of methionine using hypervalent iodine reagent **54** and thiourea.

Using phosphine reagents, a Staudinger-inspired cleavage of the methionine modification was developed, exhibiting the stimuli-responsive reversibility of the labelling strategy. A visible light-mediated reduction of the diazo moiety was also developed, which generated trialkylsulfonium products using photoredox catalysis. Stability studies were carried out on the β-sulfonium α-diazoester and reduced trialkylsulfonium conjugates, with the trialkylsulfonium products displaying significantly improved stability. The β-sulfonium α-diazoester products were amenable to isolation on a larger, milligram scale, facilitating extensive characterisation

which confirmed both the site of modification at methionine and the retention of secondary structural features following modification.

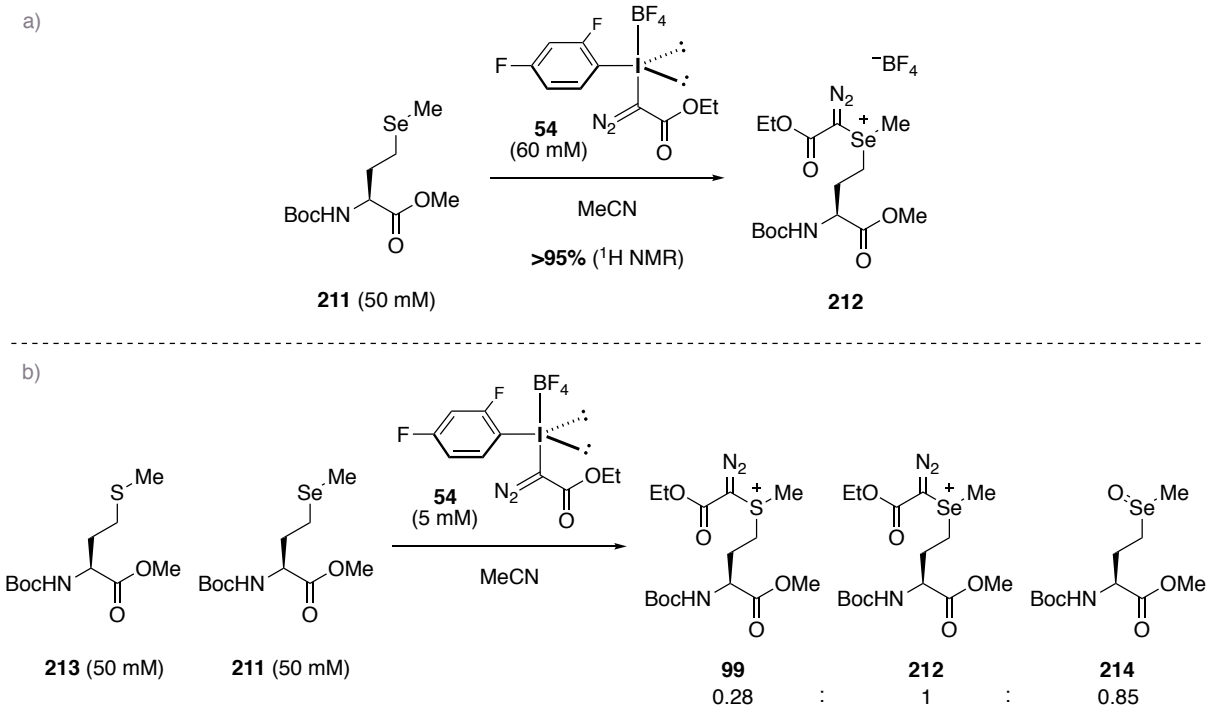
While stability studies of reduced trialkylsulfonium conjugates indicated longer half-lives compared to the β-sulfonium α-diazoester analogues, future improvements to stability would be desirable to facilitate *in vivo* application of this bioconjugation technique. It has been shown by Deming *et al.* that demethylation of alkylated methionine derivatives could be achieved to form alkylated homocysteine analogues using ammonium pyrrolidinedithiocarbamate, APDC (see Section 1.3).[188] Investigation into the demethylation of trialkylsulfonium products, formed in the two-step methionine bioconjugation and visible light-mediated reduction sequence, could provide further stability improvements for resulting conjugates (Scheme 93). Additionally, the incipient species would be neutral rather than cationic and would likely possess distinct physical properties in comparison to the sulfonium analogues, potentially rendering purification and handling more facile.



Scheme 93 Proposed demethylation of trialkylsulfonim conjugates to form neutral methionine analogues, based upon previous work by Deming.[188]

Building on the success of the visible light-mediated reduction of the diazo motif in β-sulfonium α-diazoester conjugates, a visible light-mediated benzylation was developed using C4-substituted Hantzsch ester reagents. A benzyl radical, generated from the Hantzsch ester reagent, could be formally transferred to the α-position of the β-sulfonium α-diazoester, with extrusion of dinitrogen, yielding a C-benzylation product. This functionalisation of the non-native diazo moiety provided a platform for subsequent bioorthogonal transformations of biomolecules, following methionine bioconjugation. Ongoing work within the group involves the exploration of alternative radical sources which may facilitate the transfer of more varied payloads to the diazo motif.

It has been established that the selenium analogue of methionine, selenomethionine **211**, can react with iodonium salt **54** in quantitative conversion to form cationic selenium species **212** (Scheme 94a). A competition experiment found that selenomethionine **211** can react with iodonium salt **54** in preference to methionine **213** in the ratio of 1:0.28 (alongside formation of selenoxide **214**, Scheme 94b). Selenomethionine was able to react with **54** approximately 3.5 times faster than methionine due to its increased nucleophilicity. This superior reactivity of selenomethionine could facilitate the development of an analogous functionalisation reaction, in which this naturally occurring amino acid could be incorporated in place of a methionine residue.[275] Selectivity for selenomethionine over methionine using this method may find useful application in the profiling of complex proteomes.[132]



Scheme 94 a) Reaction of selenomethionine with iodonium salt **54**. b) Competition experiment between methionine and selenomethionine with iodonium salt **54**

Considering the emergence of interest in the development of dual modification strategies for peptides and proteins, the methionine bioconjugation developed herein was applied to the synthesis of multi-functionalised polypeptides *via* combination with a tryptophan-selective bioconjugation technique. Through merging a tryptophan-selective organoradical reagent with the optimised methionine labelling conditions, a simultaneous dual-functionalisation was developed. The dual-labelling strategy was shown to be compatible with various polypeptide substrates, allowing facile synthesis of multi-functionalised scaffolds in a single step from

native biomolecules. It is hoped that the simultaneous dual-labelling will find valuable application in FRET imaging for biochemical use.[276]

5 Experimental

5.1 General Information

Proton nuclear magnetic resonance (1H NMR) spectra were recorded at ambient temperature on a Bruker AM 400 (400 MHz) or an Avance 500 (500 MHz) spectrometer. Chemical shifts (δ) are reported in ppm and quoted to the nearest 0.01 ppm relative to the residual protons in chloroform-d (7.26 ppm) or acetonitrile-d3 (1.94 ppm) and coupling constants (J) are quoted in Hertz (Hz). Data are reported as follows: Chemical shift (multiplicity, coupling constants, number of protons, assignment). Multiplicity is reported according to the following convention: s = singlet, d = doublet, t = triplet, q = quartet, qn = quintet, sext = sextet, sp = septet, m = multiplet, br = broad. Where coincident coupling constants have been observed, the apparent (app) multiplicity of the proton resonance has been reported. Carbon nuclear magnetic resonance (13C NMR) spectra were recorded at ambient temperature on a Bruker AM 400 (100 MHz) or an Avance 500 (125 MHz) spectrometer. 19F NMR spectra were recorded at 376 MHz on a Bruker AVIII spectrometer. Chemical shift (δ) was measured in ppm and quoted to the nearest 0.1 ppm relative to the residual solvent peaks in chloroform-*d* (77.16 ppm) or acetonitrile-d_3 (1.32 ppm). Infrared (IR) spectra were recorded on a Perkin Elmer 1FT-IR Spectrometer fitted with an ATR sampling accessory as either solids or neat films, either through direct application or deposited in chloroform, with absorptions reported in wavenumbers (cm^{-1}) Analytical thin layer chromatography (TLC) was performed using pre-coated Merck glass backed silica gel plates (Silicagel 60 F254). Flash column chromatography was performed using Material Harvest silica gel (230–400 mesh) under a positive pressure of compressed air unless otherwise stated. Visualization was achieved using ultraviolet light (254 nm) or with basic potassium permanganate solutions as appropriate. Tetrahydrofuran, toluene, hexane, diethyl ether and dichloromethane were dried and distilled using standard methods. For reactions requiring degassed solvents, the solvents were degassed with a positive pressure of nitrogen for a minimum of one hour before use. High-resolution mass spectrometry was performed by the University of Cambridge, Department of Chemistry Mass Spectrometry Service. MS/MS spectra were obtained by the Oncology iMED Biotech unit, AstraZeneca, Cambridge and Analytical Sciences, MedImmune, Cambridge, UK. Circular Dichroism experiments were performed on an Applied Photophysics Chirascan Circular Dichroism Spectrophotometer.

Reactions that did not involve protein modification were carried out using glassware that was dried under vacuum using a heat gun and run under an atmosphere of nitrogen unless otherwise stated. All reactions were monitored by TLC or LCMS (Shimadzu LCMS 2020; solvents: A: 0.05% Formic acid in H_2O B: 0.05% formic acid in 95:5 CH_3CN:H_2O; method: 5-95%B over 3.5 min, then hold 0.5 min, then 95-5% B over 0.5 min, hold 0.5 min, 0.7 mL/min) as appropriate. Exenatide•OAc, teriparatide•OAc, and aviptadil•OAc were purchased from Selleckchem; tetracosactide•OAc was purchased from Alpha-Diagnostics. Met-Gly, Aprotinin (from bovine lung, 3–7 TIU/mg, received containing carrying amounts of methionine sulfoxide), Ubiquitin (from bovine erythrocytes, BioUltra grade) and α-Lactalbumin (type III, >85%) were purchased from Sigma-Aldrich. All peptides and proteins were used without any additional purification. All other reagents were acquired from commercial sources and used as received. Protein modification reactions were monitored and outcomes assessed using either a Shimadzu LCMS-2020 (solvents: A: 0.05% Formic acid in H_2O B: 0.05% formic acid in 95:5 CH_3CN:H_2O) equipped with an APCI probe, or a Waters XEVO GII-S Q-TOF equipped with an ESI probe (solvent system: A: 0.1% formic acid in H_2O, B: 0.1% formic acid in 95:5 CH_3CN:H_2O). Conversions and product ratios were estimated either using total ion count (TIC, Shimadzu LCMS-2020), or MaxEnt software[7] (Waters XEVO). Note: Due to the formation of cationic sulfur, the MaxEnt software underestimates the mass of the parent protein+label peak by 1 hydrogen.

For MS/MS analysis, frozen peptide stocks were thawed on ice followed by gentle vortexing. Each peptide was diluted to a final concentration of 1 μM in 50% v/v acetonitrile, 0.1% v/v acetic acid for mass spectrometry analysis. Peptide ions were generated by electrospray ionisation (ESI) using a TriVersa NanoMate (Advion BioSciences, Inc., Ithaca, NY) from a 10 μL injection with an applied spray voltage of 1.55 kV and a gas pressure of 0.7 psi. Ions generated by ESI were introduced directly into an Orbitrap Fusion tribrid mass spectrometer (Thermo Fisher Scientific, San Jose, CA). An accurate mass full scan from 100-2000 *m/z* was performed at a resolution of 120,000 (full width at half-maximum, FWHM) with data acquisition for 1 minute using the mass spectrometer instrument control software, Orbitrap Tune Application 2.0. Peptide species charge states in the collected raw spectra were determined by manual interpretation using the theoretical monoisotopic mass. The most intense observed charge state for each peptide was selected for higher energy collisional dissociation (HCD) and electron transfer dissociation (ETD) fragmentation.

Precursor ion isolation was performed in the mass selecting quadrupole with an isolation window of 2 *m/z*. The precursor automatic gain control (AGC) target value was $1.5x10^5$ with a maximum injection time of 200 msec. The normalised collision energy (NCE) for HCD fragmentation was optimised to 30%. ETD ion reaction time was set to 100 msec with an anion AGC of 2.5e5 and a maximum injection time of 200 msec. Supplemental collisional activation (SA) in EThcD was optimised to 25%. Spectra for both fragmentation methods were acquired over a mass range of 100-4000 *m/z.* Each accurate mass full scan and subsequent MS2 fragmentation strategy was analysed in duplicate.

Peptide MS2 spectra from HCD and ETD fragmentation approaches were loaded into mMass 5.5 (http://www.mmass.org) for processing and deconvolution.[277] Observed peptide fragments were matched to theoretical fragment predictions for each peptide allowing a maximum charge state of 5 and a match tolerance of 20 ppm. The Lorikeet spectrum viewer (http://code.google.com/p/lorikeet/) was used to view combined ETD and HCD MS2 spectra annotated with y, b, c and z fragment ions. Fragmentation data were further validated by manual inspection of the spectra.

Affiliations

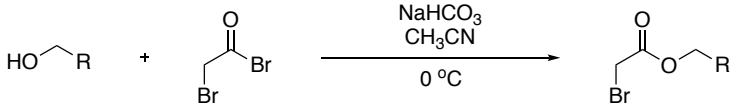
Professor Jason Chin, Dr Nicolas Huguenin and Dr Mark Skehel MRC Laboratory of Molecular Biology, Cambridge, UK

Dr Hilary J. Lewis, Bioscience, Oncology, IMED Biotech Unit, AstraZeneca, Cambridge, UK

Dr Matthew J. Edgeworth, Analytical Sciences, MedImmune, Cambridge, UK

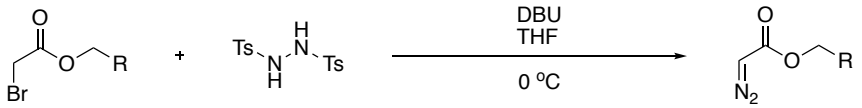
5.2 Experimental procedures for the synthesis of hypervalent iodine reagents

***General Procedure A**: Synthesis of substituted α-bromo esters*



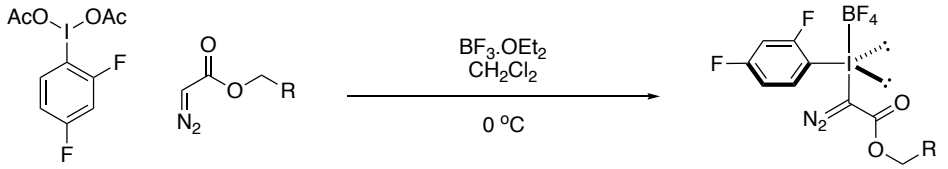
The synthesis of substituted α-bromoesters was achieved by adapting the procedure of Fukuyama.[278] To a stirring solution of sodium bicarbonate (5.0 eq.) in CH_3CN (0.15 M in alcohol), under N_2, was added the desired alcohol (1.0 eq.), and the resulting mixture cooled to 0 °C. Bromoacetyl bromide (1.5 eq.) was added, and the resulting mixture stirred for 1 - 1.5 hours at room temperature. The reaction mixture was quenched with $NaHCO_3$ (sat. aq.) and diluted with CH_2Cl_2. The organic layer was washed with $NaHCO_3$ (sat. aq.) and brine, dried over $MgSO_4$, filtered and evaporated to yield the desired α-bromoester.

***General Procedure B**: Synthesis of substituted α-diazo esters from α-bromoesters*



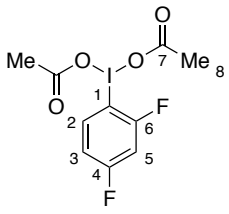
The synthesis of substituted α-diazoesters was achieved by adapting the procedure of Fukuyama.[278] To a stirring solution of *N,N'*-ditosylhydrazine (1.5 eq.) in tetrahydrofuran (0.2 M in α-bromoester) under N_2, was added the desired α-bromoester (1.0 eq.) and the resulting solution cooled to 0 °C. 1,8-diaza-[4.4.0]-bicyclo-undec-7-ene (3.0 eq.) was added dropwise and the resulting mixture stirred for 1 hour at 0 °C. The reaction was quenched with $NaHCO_3$ (sat. aq.) and diluted with specified solvent. The aqueous layer was extracted with specified solvent and the combined organic layers washed with brine, dried over $MgSO_4$, filtered and evaporated to yield the crude product. Purification by silica gel column chromatography furnished the desired α-diazoester.

***General Procedure C**: Synthesis of iodonium salts from α-diazoesters*



A vigorously stirring solution of freshly recrystallized (2,4-difluorophenyl)-λ^3-iodanediyl diacetate (1.0 eq.) in CH_2Cl_2 (0.15 M in diacetate) under N_2, was cooled to 0 °C. To the stirring solution was added $BF_3.OEt_2$ (1.1 - 1.3 eq.) and the resulting solution stirred for 5 minutes. The desired α-diazocarbonyl (1.1 - 1.3 eq.) was added rapid dropwise and the resulting solution stirred for 30 minutes at 0 °C. The solvent was evaporated, and the resulting orange/red residue was triturated using a sonicator with specified solvent to yield the desired iodonium salt.

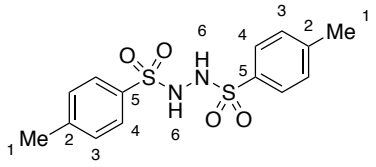
(2,4-Difluorophenyl)-λ^3-iodanediyl diacetate **76**



Sodium perborate tetrahydrate (19.3 g, 125 mmol) was added portion-wise over 20 minutes to a stirring solution of 2,4-difluoroiodobenzene (1.5 mL, 12.5 mmol) in acetic acid (110 mL) at 45 °C. Trifluoromethanesulfonic acid (6.7 mL, 75.2 mmol) was added and the resulting suspension stirred overnight at 45 °C. The solvent was removed in vacuo and the resulting slurry diluted with H_2O (50 mL). The aqueous layer was extracted with CH_2Cl_2 (3 × 60 mL), dried over $MgSO_4$, filtered and evaporated. The crude solid was recrystalised from CH_2Cl_2 and petroleum ether to yield **76** (1.99 g, 45%) as a white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 8.16 – 8.08 (m, 1H, H_2), 7.10 (app. td, J = 8.4, 2.6 Hz, 1H, H_5), 7.05 – 6.95 (m, 1H, H_3), 1.98 (s, 6H, H_8); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 177.1 (C_7), 165.97 (dd, J = 257, 11 Hz, C_4 or C_6), 160.8 (dd, J = 255, 13 Hz, C_4 or C_6), 138.5 (dd, J = 10, 3 Hz, C_2), 114.3 (dd, J = 22, 3 Hz, C_3), 105.4 (app. t, J = 26 Hz, C_5), 104.4 (dd, J = 24, 4 Hz, C_1), 20.3 (C_8); **^{19}F{^{1}H}-NMR** (376 MHz, CD_3CN) δ: -91.4 (d, J = 12 Hz), -100.7 (d, J = 12 Hz)

*Data in accordance with literature[279]

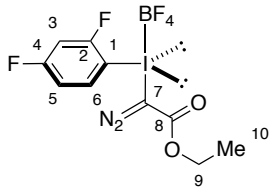
N,N'-Ditosylhydrazine **70**



p-Toluenesulfonyl chloride (8.3 g, 43.5 mmol) was added to a solution of p-toluenesulfonyl hydrazide (6.2 g, 33.5 mmol) in CH_2Cl_2 (34 mL), under N_2, and the resulting solution cooled to 0 °C. Pyridine (3.5 mL, 43.5 mmol) was added dropwise over 5 minutes and the reaction mixture turned yellow and became homogeneous. The reaction mixture was stirred at 0 °C for 30 minutes, then n-hexane (50 mL) and H_2O (75 mL) were added and the mixture stirred for a further 15 minutes at 0 °C. The resultant suspension was filtered and the crude solid recrystalised from acetone (80 mL) and H_2O (38 mL). The precipitates were collected by filtration, washed with ice-cooled Et_2O (25 mL) and dried in vacuo to yield **70** (9.8 g, 86%) as a white solid. **^{1}H-NMR** (400 MHz, $(CD_3)_2SO$) δ: 9.58 (br. s, 2H, H_6), 7.64 (d, J = 8.3 Hz, 4H, H_4), 7.38 (d, J = 8.3 Hz, 4H, H_3), 2.39 (s, 6H, H_1); **^{13}C-NMR** (100 MHz, $(CD_3)_2SO$) δ: 143.5 (C_2), 135.6 (C_5), 129.5 (C_3), 127.8 (C_4), 21.1 (C_1)

*Data in accordance with literature[278]

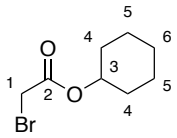
(1-Diazo-2-ethoxy-2-oxoethyl)(2,4-difluorophenyl)iodonium tetrafluoroborate **54**



54 was prepared according to General Procedure C using (2,4-difluorophenyl)-λ^3-iodanediyl diacetate (1.0 g, 2.8 mmol), boron trifluoride diethyl etherate (0.45 mL, 3.6 mmol), and ethyl diazoacetate (0.46 mL (15% wt. CH_2Cl_2), 3.6 mmol), with a reaction time of 30 minutes. Sequential trituration using a sonicator (1:9 CH_2Cl_2:diethyl ether (5 × 20 mL), 40-60 petroleum (5 × 10 mL), diethyl ether (3 × 10 mL)) afforded **54** (1.0 g, 81%) as a yellow, free-flowing powder. **54** shows good stability (>3 months) when stored in a –20 °C freezer. **54** displays the following stability profile in D_2O: 24 hours: 90% remaining, 48 hours: 69% remaining. **^{1}H-**

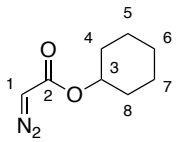
NMR (400 MHz, CD_3CN) δ: 8.26 – 8.18 (m, 1H, H_6), 7.36 (app. td, J = 8.8, 2.7 Hz, 1H, H_3), 7.25 – 7.19 (m, 1H, H_5), 4.28 (q, J = 7.1 Hz, 2H, H_9), 1.26 (t, J = 7.2 Hz, 3H, H_{10}); **^{13}C-NMR** (125 MHz, CD_3CN) δ: 168.0 (dd, J = 257, 12 Hz, C_2 or C_4), 162.5 (dd, J = 257, 12 Hz, C_2 or C_4), 161.5 (C_8), 140.1 (dd, J = 12, 2 Hz, C_6), 116.4 (dd, J = 23, 3 Hz, C_5), 107.5 – 106.9 (m, C_3), 97.6 (dd, J = 24, 4 Hz, C_1), 65.4 (C_9), 42.4 (br. s, C_7), 14.5 (C_{10}); **^{19}F{^{1}H}-NMR** (376 MHz, CD_3CN) δ: -90.8 (d, J = 12 Hz), -96.8 (d, J = 12 Hz), -144.1 (s), -144.2 (m); **^{1}H{^{19}F}-NMR** (400 MHz, CD_3CN) δ: 8.20 (d, J = 9.0 Hz, 1H, H_6), 7.37 (d, J = 2.8 Hz, 1H, H_3), 7.22 (dd, J = 9.0, 2.8 Hz, 1H, H_5), 4.28 (q, J = 7.1 Hz, 2H, H_9), 1.26 (t, J = 7.1 Hz, 3H, H_{10}); **IR** ν_{max}/cm^{-1} (film): 2989, 2117, 1682; **HRMS**: $(ESI)^+$ (m/z) $[M - BF_4]^+$: Found 352.9590, Calc. $C_{10}H_8F_2IN_2O_2^+$ requires 352.9593; **m.p.**: 64-66 °C

Cyclohexyl 2-bromoacetate **65**



65 was prepared according to General Procedure A using cyclohexanol (0.21 mL, 2.0 mmol), $NaHCO_3$ (0.50 g, 6.0 mmol) and bromoacetyl bromide (0.25 mL, 2.8 mmol) in CH_3CN (10 mL), with a reaction time of 2 hours. The reaction mixture was diluted with CH_2Cl_2 (20 mL) and $NaHCO_3$ (sat. aq.) (10 mL). The organic layer was washed with $NaHCO_3$ (sat. aq.) (2 × 10 mL) and brine (10 mL), dried over $MgSO_4$, filtered and evaporated to yield **65** (0.43 g, 98%) which was used directly without further purification. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 4.80 (app. sp, J = 3.9 Hz, 1H), 3.79 (s, 2H), 1.89 – 1.80 (m, 2H), 1.77 – 1.67 (m, 2H), 1.57 – 1.23 (m, 6H)

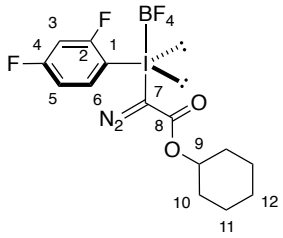
Cyclohexyl 2-diazoacetate **71**



71 was prepared according to General Procedure B using **65** (0.22 g, 1.0 mmol), N,N'-ditosylhydrazine **70** (0.50 g, 1.5 mmol), and 1,8-diazo-bicyclo-[5.4.0]-undec-7-ene (0.41 mL,

3.0 mmol), with a reaction time of 1 hour. Column chromatography (10% diethyl ether/40-60 petroleum ether) afforded **71** (0.12 g, 72%) as a yellow oil. **^{1}H-NMR** (500 MHz, $CDCl_3$) δ: 4.81 (app. sp, J = 4.0 Hz, 1H, H_3), 4.69 (br. s, 1H, H_1), 1.90 – 1.78 (m, 2H, H_4' and H_8'), 1.76 – 1.63 (m, 2H, H_5' and H_7'), 1.56 – 1.46 (m, 1H, H_6'), 1.45 – 1.29 (m, 4H, H_4'', H_5'', H_7'', H_8''), 1.28 – 1.16 (m, 1H, H_6''); **^{13}C-NMR** (125 MHz, $CDCl_3$) δ: 166.4 (C_2), 73.3 (C_3), 46.3 (C_1), 31.9 (C_4 and C_8), 25.4 (C_6), 23.7 (C_5 and C_7); **IR** ν_{max}/cm^{-1} (film): 2935, 2860, 2102, 1677, 1181; **HRMS**: $(ESI)^+$ (m/z) $[M+H]^+$: Found 169.0971, Calc. $C_{10}H_8F_2IN_2O_2^+$ requires 169.0972

(2-(Cyclohexyloxy)-1-diazo-2-oxoethyl)(2,4-difluorophenyl)iodonium tetrafluoroborate **77**

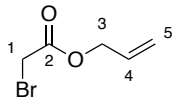


77 was prepared according to General Procedure C using (2,4-difluorophenyl)-λ^3-iodanediyl diacetate **76** (0.36 g, 1.0 mmol), boron trifluoride diethyl etherate (0.18 mL, 1.3 mmol), and cyclohexyl diazoacetate **71** (0.16 g, 1.3 mmol), with a reaction time of 30 minutes. Trituration using a sonicator (9:1 diethyl ether:40-60 petroleum ether, 20 mL, followed by 1:1 diethyl ether:40-60 petroleum ether, 20 mL) afforded **77** (0.27 g, 66%) as a yellow, free-flowing, powder. **^{1}H-NMR** (500 MHz, CD_3CN) δ: 8.22 – 8.15 (m, 1H, H_6), 7.37 (app. td, J = 8.6, 2.6 Hz, 1H, H_3), 7.26 – 7.19 (m, 1H, H_5), 4.86 (app. sp, J = 3.9 Hz, 1H, H_9), 1.87 – 1.78 (m, 2H, H_{10}'), 1.75 – 1.65 (m, 2H, H_{11}'), 1.55 – 1.26 (m, 6H, H_{10}'', H_{11}'', H_{12}); **^{13}C-NMR** (125 MHz, CD_3CN) δ: 167.9 (dd, J = 257, 11 Hz, C_2 or C_4), 162.5 (dd, J = 256, 13 Hz, C_2 or C_4), 160.7 (C_8), 139.9 (dd, J = 11, 2 Hz, C_6), 116.4 (dd, J = 23, 3 Hz, C_5), 107.4 – 106.9 (m, C_3), 97.6 (dd, J = 23, 4 Hz, C_1), 78.3 (C_9), 31.9 (C_{10}'), 25.7 (C_{12}), 23.9 (C_{11}')*; **^{19}F{^{1}H}-NMR** (376 MHz, CD_3CN) δ: -90.7 (d, J = 12 Hz), -96.9 (d, J = 12 Hz), -144.1 (s), -144.2 (s); **IR** ν_{max}/cm^{-1} (film): 2951, 2114, 1698, 1034; **HRMS**: $(ESI)^+$ (m/z) $[M - BF_4]^+$: Found 407.0068, Calc. $C_{14}H_{14}F_2IN_2O_2^+$ requires 407.0063; **m.p.**: 78-80 °C (decomp.)

***** C_7 not observed

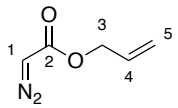
Notes: additional precipitate can be collected. While the carbon possessing the diazo moiety was not observable by ^{13}C-NMR, the IR stretch[8] at 2114 cm^{-1} as well as the correct mass being observed confirms the presence of the diazo group.

Allyl 2-bromoacetate **66**



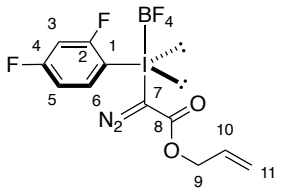
66 was prepared according to General Procedure A using allyl alcohol (0.55 mL, 8.0 mmol), $NaHCO_3$ (2.02 g, 24.0 mmol) and bromoacetyl bromide (0.98 mL, 11.2 mmol) in CH_3CN (40 mL), with a reaction time of 2 hours. The reaction mixture was diluted with CH_2Cl_2 (20 mL) and washed with $NaHCO_3$ (sat. aq.) (3 × 10 mL) and brine (10 mL), dried over $MgSO_4$, filtered and evaporated to yield the **66** (1.17 g, 82%) which was used directly without further purification. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 5.97 – 5.86 (m, 1H), 5.40 – 5.24 (m, 2H), 4.68 – 4.63 (m, 2H), 3.86 (br. s, 2H)

Allyl 2-diazoacetate **72**



72 was prepared according to General Procedure B using **66** (1.2 g, 6.5 mmol), N,N'-ditosylhydrazine (3.3 g, 9.8 mmol), and 1,8-diazo-bicyclo-[5.4.0]-undec-7-ene (3.6 mL, 26.0 mmol), with a reaction time of 1 hour. Diethyl ether (3 × 50 mL) was used in the workup. Column chromatography (10-30% diethyl ether/40-60 petroleum ether) afforded **72** (0.5 g, 61%) as a volatile yellow oil which was used directly. **^{1}H-NMR** (500 MHz, $CDCl_3$) δ: 6.00 – 5.89 (m, 1H), 5.37 – 5.19 (m, 2H), 4.76 (br. s, 1H), 4.67 – 4.63 (m, 2H)

Allyl 2-diazo-2-((2,4-difluorophenyl)(tetrafluoro-λ^5-boraneyl)-λ^3-iodaneyl)acetate **78**

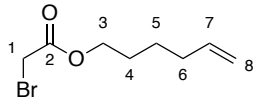


78 was prepared according to General Procedure C using (2,4-difluorophenyl)-λ^3-iodanediyl diacetate **76** (1.1 g, 3.1 mmol), boron trifluoride diethyl etherate (0.5 mL, 4.0 mmol), and allyl diazoacetate **72** (0.50 g, 4.0 mmol), with a reaction time of 30 minutes. Trituration using a sonicator (9:1 diethyl ether:40-60 petroleum ether, 20 mL, followed by 3:1 diethyl ether:40-60 petroleum ether, 20 mL) afforded **78** (0.68 g, 60%) as a yellow oil. **^{1}H-NMR** (500 MHz, CD_3CN) δ: 8.25 – 8.18 (m, 1H, H_6), 7.35 (app. td, J = 8.7, 2.6 Hz, H_3), 7.26 – 7.18 (m, 2H, H_5), 5.98 – 5.87 (m, 1H, H_{10}), 5.39 – 5.25 (m, 2H, H_{11}), 4.73 (app. d, J = 5.7 Hz, 2H, H_9); **^{13}C-NMR** (125 MHz, CD_3CN) δ: 168.0 (dd, J = 257, 12 Hz, C_2 or C_4), 162.5 (dd, J = 254, 13 Hz, C_2 or C_4), 161.3 (C_8), 140.1 (dd, J = 11, 1 Hz, C_6), 132.3 (C_{10}), 119.8 (C_{11}), 116.5 (dd, J = 24, 3 Hz, C_5), 107.6 – 106.9 (m, C_3), 97.7 (dd, J = 24, 4 Hz, C_1), 69.2 (C_9)*; **^{19}F{^{1}H}-NMR** (376 MHz, CD_3CN) δ: -92.4 (d, J = 12 Hz), -99.7 (d, J = 12 Hz), -148.8 (s), -150.5 (s); **IR** ν_{max}/cm^{-1} (film): 2988, 2118, 1701, 1036; **HRMS**: (ESI)$^+$ (m/z) [M – BF_4]$^+$: Found 364.9591, Calc. $C_{11}H_8F_2IN_2O_2^+$ requires 364.9593

***** C_7 not observed

Notes: While the carbon possessing the diazo moiety was not observable by ^{13}C-NMR, the IR stretch[8] at 2118 cm^{-1} as well as the correct mass being observed confirms the presence of the diazo group.

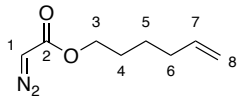
Hex-5-en-1-yl 2-bromoacetate **67**



67 was prepared according to General Procedure A using 5-hexen-1-ol (0.24 mL, 2.0 mmol), $NaHCO_3$ (0.50 g, 6.0 mmol) and bromoacetyl bromide (0.25 mL, 2.8 mmol) in CH_3CN (10 mL), with a reaction time of 2 hours. The reaction mixture was diluted with CH_2Cl_2 (20 mL)

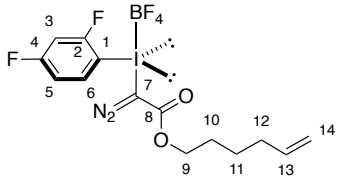
and washed with $NaHCO_3$ (sat. aq.) (3 × 10 mL) and brine (10 mL), dried over $MgSO_4$, filtered and evaporated to yield the **67** (0.24 g, 56%) which was used directly without further purification. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 5.81 – 5.70 (m, 1H), 5.02 – 4.91 (m, 2H), 4.15 (t, J = 6.2 Hz, 2H), 3.80 (s, 2H), 2.11 – 2.01 (m, 2H), 1.71 – 1.61 (m, 2H), 1.50 – 1.39 (m, 2H)

Hex-5-en-1-yl 2-diazoacetate **73**



73 was prepared according to General Procedure B using **76** (0.25 g, 1.1 mmol), N,N'-ditosylhydrazine **70** (0.57 g, 1.7 mmol), and 1,8-diazo-bicyclo-[5.4.0]-undec-7-ene (0.46 mL, 3.3 mmol), with a reaction time of 1 hour. Column chromatography (10% diethyl ether/40-60 petroleum ether) afforded **73** (0.09 g, 48%) as a yellow oil. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 5.83 – 5.71 (m, 1H, H_7), 5.03 – 4.91 (m, 2H, H_8), 4.72 (br. s, 1H, H_1), 4.14 (t, J = 6.8 Hz, 2H, H_3), 2.10 – 2.02 (m, 2H, H_6), 1.68 – 1.59 (m, 2H, H_4), 1.48 – 1.38 (m, 2H, H_5); **^{13}C-NMR** (125 MHz, $CDCl_3$) δ: 166.9 (C_2), 138.4 (C_7), 115.0 (C_8), 64.8 (C_3), 46.2 (C_1), 33.4 (C_6), 28.3 (C_4), 25.2 (C_5); **IR** ν_{max}/cm^{-1} (film): 2935, 2105, 1685, 1179; **HRMS**: $(ESI)^+$ (m/z) $[M+H]^+$: Found 169.0971, Calc. $C_8H_{13}O_2N_2^+$ requires 169.0972

Hex-5-en-1-yl 2-diazo-2-((2,4-difluorophenyl)(tetrafluoro-λ^5-boraneyl)-λ^3-iodaneyl)acetate **79**



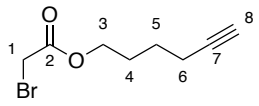
79 was prepared according to General Procedure C using (2,4-difluorophenyl)-λ^3-iodanediyl diacetate **76** (1.15 g, 3.2 mmol), boron trifluoride diethyl etherate (0.51 mL, 4.2 mmol), and hexenyl diazoacetate **73** (0.70 g, 4.1 mmol), with a reaction time of 30 minutes. Trituration using a sonicator (9:1 diethyl ether:40-60 petroleum ether, 40 mL, followed by 3:1 diethyl ether:40-60 petroleum ether, 20 mL) afforded **79** (1.14 g, 88%) as a yellow oil. **^{1}H-NMR** (500 MHz, CD_3CN) δ: 8.24 – 8.15 (m, 1H, H_6), 7.37 (app. td, J = 8.8, 2.8 Hz, 1H, H_3),7.26 – 7.18

(m, 1H, H_5), 5.87 – 5.74 (m, 1H, H_{13}), 5.09 – 4.90 (m, 2H, H_{14}), 4.24 (t, J = 6.5 Hz, 2H, H_9), 2.10 – 2.02 (m, 2H, H_{12}), 1.68 – 1.62 (m, 2H, H_{10}), 1.44 – 1.37 (m, 2H, H_{11}); **^{13}C-NMR** (125 MHz, $CDCl_3$) δ: 168.1 (dd, J = 257, 12 Hz, C_2 or C_4), 162.6 (dd, J = 254, 12 Hz, C_2 or C_4), 161.5 (C_8), 140.0 (app. d, J = 11 Hz, C_6), 139.5 (C_{13}), 116.5 (dd, J = 23, 3 Hz, C_5), 115.3 (C_{14}), 107.6 – 106.9 (m, C_3), 97.7 (dd, J = 24, 4 Hz, C_1), 69.1 (C_9), 33.8 (C_{12}), 28.7 (C_{10}), 25.6 (C_{11})*; **^{19}F{^{1}H}-NMR** (376 MHz, CD_3CN) δ: -92.6 (d, J = 12 Hz), -99.7 (d, J = 12 Hz), -150.7 (s), -150.8 (s); **IR** ν_{max}/cm^{-1} (film): 2952, 2117, 1702, 1031; **HRMS**: $(ESI)^+$ (m/z) $[M - BF_4]^+$: Found 407.0067, Calc. $C_{11}H_8F_2IN_2O_2^+$ requires 407.0063

***** C_7 not observed

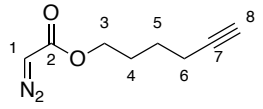
Notes: While the carbon possessing the diazo moiety was not observable by ^{13}C-NMR, the IR stretch[8] at 2117 cm^{-1} as well as the correct mass being observed confirms the presence of the diazo group.

Hex-5-yn-1-yl 2-bromoacetate **68**



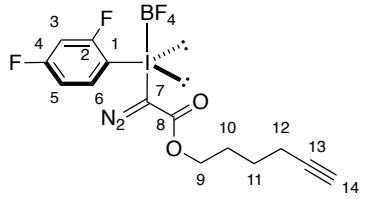
68 was prepared according to General Procedure A using 5-hexyn-1-ol (0.22 mL, 2.0 mmol), $NaHCO_3$ (0.50 g, 6.0 mmol) and bromoacetyl bromide (0.25 mL, 2.8 mmol) in CH_3CN (10 mL), with a reaction time of 2 hours. The reaction mixture was diluted with CH_2Cl_2 (20 mL) and washed with $NaHCO_3$ (sat. aq.) (3 × 10 mL) and brine (10 mL), dried over $MgSO_4$, filtered and evaporated to yield **68** (0.39 g, 89%) which was used directly without further purification. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 4.18 (t, J = 6.6 Hz, 2H), 3.81 (s, 2H), 2.22 (td, J = 7.0, 2.7 Hz, 2H), 1.95 (t, J = 2.7 Hz, 1H), 1.82 – 1.74 (m, 2H), 1.64 – 1.55 (m, 2H)

Hex-5-yn-1-yl 2-diazoacetate **74**



74 was prepared according to General Procedure B using **68** (0.23 g, 1.0 mmol), N,N'-ditosylhydrazine **70** (0.53 g, 1.5 mmol), and 1,8-diazo-bicyclo-[5.4.0]-undec-7-ene (0.41 mL, 3.0 mmol), with a reaction time of 1 hour. Column chromatography (10% diethyl ether/40-60 petroleum ether) afforded **74** (0.12 g, 69%) as a yellow oil. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 4.71 (br. s, 1H, H_1), 4.12 (t, J = 7.2 Hz, 2H, H_3), 2.17 (td, J = 7.2, 2.7 Hz, 2H, H_6), 1.92 (t, J = 2.7 Hz, 1H, H_8), 1.76 – 1.67 (m, 2H, H_4), 1.58 – 1.49 (m, 2H, H_5); **^{13}C-NMR** (125 MHz, $CDCl_3$) δ: 167.1 (C_2), 83.7 (C_7), 68.8 (C_8), 64.3 (C_3), 46.1 (br. s, C_1), 27.8 (C_4), 24.8 (C_5), 18.0 (C_6); **IR** ν_{max}/cm^{-1} (film): 2953, 2105, 1682, 1179; **HRMS**: $(ESI)^+$ (m/z) $[M+H]^+$: Found 167.0817, Calc. $C_8H_{11}N_2O_2^+$ requires 167.0815

Hex-5-yn-1-yl 2-diazo-2-((2,4-difluorophenyl)(tetrafluoro-λ^5-boraneyl)-λ^3-iodaneyl)acetate **80**



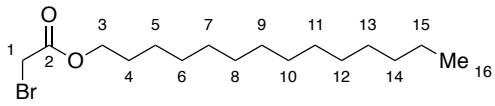
80 was prepared according to General Procedure C using (2,4-difluorophenyl)-λ^3-iodanediyl diacetate **76** (1.0 g, 2.8 mmol), boron trifluoride diethyl etherate (0.45 mL, 3.7 mmol), and hexynyl diazoacetate **74** (0.61 g, 3.7 mmol), with a reaction time of 30 minutes. Trituration using a sonicator (9:1 diethyl ether:40-60 petroleum ether, 40 mL, followed by 3:1 diethyl ether:40-60 petroleum ether, 20 mL) afforded **80** (1.3 g, 87%) as a yellow oil. **^{1}H-NMR** (500 MHz, CD_3CN) δ: 8.25 – 8.18 (m, 1H, H_6), 7.36 (app. td, J = 8.7, 2.7 Hz, H_3), 7.25 – 7.19 (m, 1H, H_5), 4.25 (t, J = 6.3 Hz, 2H, H_9), 2.23 – 2.17 (m, 3H, H_{12} and H_{14}), 1.79 – 1.69 (m, 2H, H_{10}), 1.57 – 1.48 (m, 2H, H_{11}); **^{13}C-NMR** (125 MHz, $CDCl_3$) δ: 168.0 (dd, J = 256, 11 Hz, C_2 or C_4), 162.5 (dd, J = 256, 13 Hz, C_2 or C_4), 161.5 (C_8), 140.0 (dd, J = 11, 2 Hz, C_6), 116.5 (dd, J = 23, 3 Hz, C_5), 107.5 – 106.9 (m, C_3), 97.7 (dd, J = 23, 4 Hz, C_1), 84.9 (C_{13}), 70.2 (C_{14}), 68.7 (C_9), 28.3 (C_{10}), 25.5 (C_{11}), 18.3 (C_{12})*; **^{19}F{^{1}H}-NMR** (376 MHz, CD_3CN) δ: -92.5 (d, J = 12

Hz), -99.7 (d, J = 12 Hz), -150.5 (s), -150.6 (s); **IR** ν_{max}/cm^{-1} (film): 3282, 2118, 1702, 1029; **HRMS**: $(ESI)^+$ (m/z) $[M - BF_4]^+$: Found 404.9906, Calc. $C_{14}H_{12}F_2IN_2O_2^+$ requires 404.9906

***** C_7 not observed

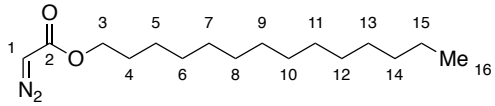
Notes: While the carbon possessing the diazo moiety was not observable by ^{13}C-NMR, the IR stretch[8] at 2118 cm^{-1} as well as the correct mass being observed confirms the presence of the diazo group.

Tetradecyl 2-bromoacetate **69**



69 was prepared according to General Procedure A using 1-tetradecanol (0.43 mL, 2.0 mmol), $NaHCO_3$ (0.50 g, 6.0 mmol) and bromoacetyl bromide (0.25 mL, 2.8 mmol) in CH_2Cl_2 (10 mL), with a reaction time of 2 hours. The reaction mixture was diluted with CH_2Cl_2 (20 mL) and washed with $NaHCO_3$ (sat. aq.) (3 × 10 mL) and brine (10 mL), dried over $MgSO_4$, filtered and evaporated to yield the **69** (0.61 g, 91%) which was used directly without further purification. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 4.16 (t, J = 6.7 Hz, 2H), 3.82 (s, 2H), 1.69 – 1.61 (m, 2H), 1.35 – 1.21 (m, 22H), 0.90 – 0.85 (m, 3H).

Tetradecyl 2-diazoacetate **75**



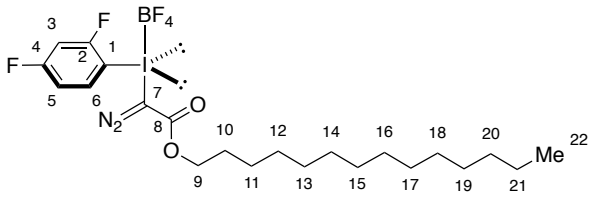
75 was prepared according to General Procedure B using **69** (0.30 g, 0.9 mmol), N,N'-ditosylhydrazine (0.46 g, 1.4 mmol), and 1,8-diazo-bicyclo-[5.4.0]-undec-7-ene (0.37 mL, 2.7 mmol), with a reaction time of 1 hour, affording **75** (0.19 g, 74%) as a yellow oil which was used without further purification. **^{1}H-NMR** (500 MHz, CD_3CN) δ: 4.72 (br. s, 1H, H_1), 4.14 (t, J = 6.7 Hz, 2H, H_3), 1.68 – 1.58 (m, 2H, H_4), 1.38 – 1.22 (m, 22H, H_5-H_{15}), 0.90 – 0.84 (m, 3H, H_{16}); **^{13}C-NMR** (125 MHz, $CDCl_3$) δ: 65.2 (C_3), 46.2 (br. s, C_1), 32.1 (*CH_2*), 29.82 (*CH_2*), 29.79

(*CH₂*), 29.78 (*CH₂*), 29.77 (*CH₂*), 29.7 (*CH₂*), 29.6 (*CH₂*), 29.5 (*CH₂*), 29.4 (*CH₂*), 28.9 (C_4), 25.9 (*CH₂*), 22.8 (*CH₂*), 14.2 (C_{16})*; **IR** ν_{max}/cm^{-1} (film): 2922, 2853, 2107, 1694, 1181; **HRMS**: $(ESI)^+$ (m/z) $[M+H]^+$: Found 283.2375, Calc. $C_{16}H_{31}N_2O_2^+$ requires 283.2386

* C_2 not observed

Notes: While the carbon possessing the diazo moiety was not observable by ^{13}C-NMR, the IR stretch[8] at 2107 cm^{-1} as well as the correct mass being observed confirms the presence of the diazo group.

Tetradecyl 2-diazo-2-((2,4-difluorophenyl)(tetrafluoro-λ⁵-boraneyl)-λ³-iodaneyl)acetate **81**



81 was prepared according to General Procedure C using (2,4-difluorophenyl)-λ³-iodanediyl diacetate **76** (0.50 g, 1.4 mmol), boron trifluoride diethyl etherate (0.22 mL, 1.8 mmol), and **75** (0.51 g, 1.8 mmol), with a reaction time of 30 minutes. The reaction mixture was diluted with water (5 mL) and extracted with CH_2Cl_2 (4 × 5 mL). The combined organic extracts were filtered through cotton and concentrated *in vacuo*. The resultant orange oil was dissolved in 40-60 petroleum (15 mL) and cooled to between -30 °C and -40 °C, resulting in the formation of a yellow oil. The solvent was decanted and the resulting oil dried under vacuum to afford **81** (0.52 g, 70%) as a yellow oil. **^{1}H-NMR** (500 MHz, CD_3CN) δ: 8.24 – 8.18 (m, 1H, H_6), 7.37 (app. td, J = 8.8, 2.8 Hz, 1H, H_3), 7.25 – 7.20 (m, 1H, H_5), 4.23 (t, J = 6.6 Hz, 2H, H_9), 1.67 – 1.60 (m, 2H, H_{10}), 1.33 – 1.24 (m, 22H, H_{11}, H_{12}, H_{13}, H_{14}, H_{15}, H_{16}, H_{17}, H_{18}, H_{19}, H_{20}, H_{21}), 0.88 (app. t, J = 6.6 Hz, 3H, H_{22}); **^{13}C-NMR** (125 MHz, $CDCl_3$) δ: 168.0 (dd, J = 257, 12 Hz, C_2 or C_4), 162.4 (dd, J = 255, 13 Hz, C_2 or C_4), 161.4 (br. s, C_8), 140.0 (dd, J = 11, 2 Hz, C_6), 116.5 (dd, J = 23, 3 Hz, C_5), 107.5 – 107.0 (m, C_3), 97.7 (dd, J = 24, 4 Hz, C_1), 69.3 (C_9), 32.7 (*CH₂*), 30.4 (*CH₂*), 30.38 (*CH₂*), 30.36 (*CH₂*), 30.35 (*CH₂*), 30.27 (*CH₂*), 30.21 (*CH₂*), 30.1 (*CH₂*), 29.8 (*CH₂*), 29.2 (C_{10}), 26.3 (*CH₂*), 23.4 (*CH₂*), 14.5 (C_{22})*; **^{19}F{^{1}H}-NMR** (376 MHz, CD_3CN) δ: -92.5 (d, J = 12 Hz), -99.7 (d, J = 12 Hz), -150.9 (s), -151.0 (s); **IR** ν_{max}/cm^{-1} (film): 2923, 2853,

2117, 1706, 1051; **HRMS**: $(ESI)^+$ (m/z) $[M - BF_4]^+$: Found 521.1475, Calc. $C_{22}H_{32}F_2IN_2O_2^+$ requires 521.1471

***** C_7 not observed

Notes: While the carbon possessing the diazo moiety was not observable by ^{13}C-NMR, the IR stretch[8] at 2117 cm^{-1} as well as the correct mass being observed confirms the presence of the diazo group.

5.3 Experimental procedures for polypeptide and protein functionalisation

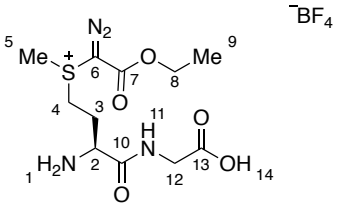
***General Procedure D**: Polypeptide and Protein Functionalization*

A 2 mL vial equipped with a magnetic stirrer was charged with a solution containing the desired protein (2 mM in 0.1 M thiourea (aq.), 20 μL). To the vial was added thiourea (100 mM in H_2O, 20 μL) and H_2O (10 μL). To the solution was added TEMPO (40 mM in H_2O, 50 μL) and the vial was stirred vigorously. To the stirring solution was added the iodonium salt (100 mM in CH_3CN or H_2O, 100 μL) and the resulting solution stirred for 5 minutes at room temperature. The resulting mixture was then washed with diethyl ether or ethyl acetate, the organic layer was decanted and discarded while the remaining organic volatiles were then removed from the aqueous layer *in vacuo*. The resulting solution was analyzed directly *via* LC/MS.

***General Procedure E**: Polypeptide and Protein Functionalization*

A 2 mL vial equipped with a magnetic stirrer was charged with a solution containing the desired protein (2 mM in 0.1 M thiourea (aq.), 20 μL). To the vial was added thiourea (100 mM in H_2O, 20 μL) and formic acid (100 mM in H_2O, 10 μL). To the solution was added TEMPO (40 mM in H_2O, 50 μL) and the vial was stirred vigorously. To the stirring solution was added the iodonium salt (100 mM in CH_3CN or H_2O, 100 μL) and the resulting solution stirred for 5 minutes at room temperature. The resulting mixture was then washed with diethyl ether or ethyl acetate, the organic layer was decanted and discarded while the remaining organic volatiles were then removed from the aqueous layer *in vacuo*. The resulting solution was analyzed directly *via* LC/MS.

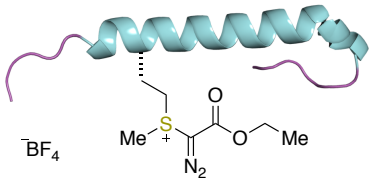
((S)-3-amino-4-((carboxymethyl)amino)-4-oxobutyl)(1-diazo-2-ethoxy-2-oxoethyl) (methyl)sulfonium tetrafluoroborate **52**



A round-bottomed flask equipped with a magnetic stir-bar was charged with methionine-glycine **51** (52 mg, 0.25 mmol) and H_2O (0.5 mL) then evacuated/backfilled with N_2 (3 cycles). To the flask was added a freshly degassed solution of **54** (164 mg in H_2O (1.0 mL)), and the resulting solution stirred for 1 hour at room temperature. The resulting suspension was then diluted with water and ethyl acetate the organic layer decanted. The aqueous layer was washed twice more with ethyl acetate, then evaporated to dryness in vacuo. The resulting yellow residue was then redissolved in a minimal amount of acetonitrile and precipitated with ether at 4 °C. The supernatant was decanted, and the residue was dried in vacuo to yield **52** as a faint yellow residue (62 mg, 78%, 1:1 mixture of diastereomers). **52** was considerably hygroscopic and was stored at –20 °C. **^{1}H-NMR** (500 MHz, CD_3CN) δ: 4.42 (q, J = 7.2 Hz, 2H, H_8), 4.31 – 4.20 (m, 1H, H_2), 4.07 – 3.90 (m, 3H, H_4 and H_{12}), 3.85 – 3.67 (m, 1H, H_4'), 3.33 (app. d, J = 3.2 Hz, 3H, H_5, H_5'), 2.60 – 2.40 (m, 2H, H_3 and H_3'), 1.34 (t, J = 7.2 Hz, 3H, H_9); **^{13}C-NMR** (125 MHz, $CDCl_3$) δ: 174.3 (C_{13}, C_{13}'), 167.9 (C_{10}), 167.8 (C_{10}'), 161.3 (C_7), 161.2 (C_7'), 64.7 (C_8), 53.8 (2C_6), 51.4 (C_2), 51.3 (C_2'), 42.4 (2C_{12}), 39.5 (C_4), 39.4 (C_4'), 26.5 (C_5, C_5'), 25.6 (C_3), 25.5 (C_3'), 13.3 (C_9); **IR** ν_{max}/cm^{-1} (film): 2988, 2150, 1684, 1005; **HRMS**: $(ESI)^+$ (m/z) $[M]^+$: Found 319.1069, Calc. $C_{11}H_{19}O_5N_4{}^{32}S_1{}^+$ requires 319.1071

5.3.1 Substrate scope for polypeptide and protein functionalisation

Exenatide acetate functionalization with **54** *in aqueous conditions forming* **56**



The labelling of exenatide was performed at room temperature using General Procedure D using **54** (100 mM in H_2O) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) for the extraction. The resulting solution was then analysed directly *via* LC/MS C18 column (50 × 2.1 mm, 2.6µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min) and was observed to have proceeded in >95% conversion.

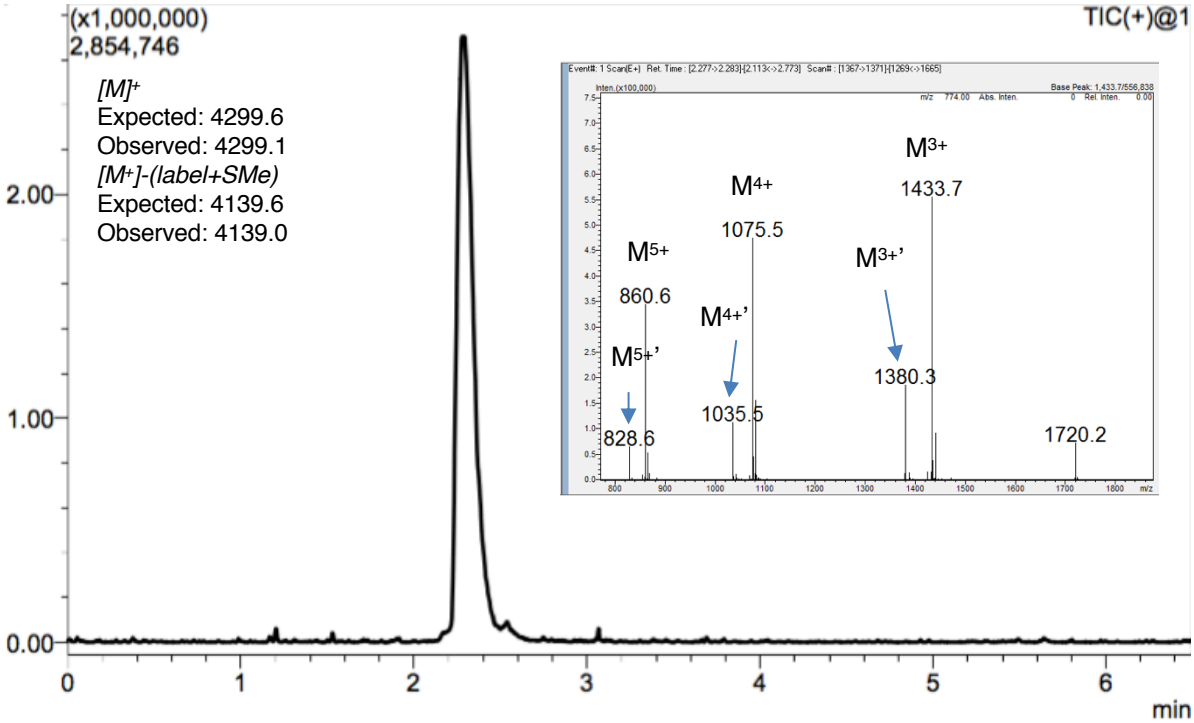


Figure 18 Mass trace of crude reaction mixture of exenatide with 54, with mass chromatogram overlayed. ' denotes diagnostic fragmentation of C–S bond giving *M+label-(label+SMe) = M+113-160.*

Large Scale Synthesis of **56**

A 2 mL vial equipped with a magnetic stirrer was charged with Exenatide•OAc (~1.8 mg, ~0.6 µmol). To the vial was added thiourea (100 mM in H_2O, 160 µL), H_2O (120 µL) and TEMPO (67 mM in H_2O, 120 µL) and the vial was stirred vigorously before **54** (100 mM in H_2O, 400 µL) was added and the resulting solution stirred for 5 minutes at room temperature. The reaction mixture was then extracted with 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL). The remaining organic volatiles were then removed in vacuo. Purification via reverse-phase HPLC* (C18 column (150 × 10 mm, 10 µm), Gradient: 5-55% B over 9 min then 55-95% B over 1 min, hold for 1.5 min, then 95-5%B over 0.5 min, hold for 3 min, 5 ml/min), followed by lyophilization yielded **56** as a sticky, off-white solid that was immediately redissolved in H_2O (500 µL) and stored at –20 °C. Concentration analysis by A_{280} indicated a 1.4 mg/mL solution ~79% isolated yield).

***Note**: Purification of **56** should be carried out expediently as prolonged purification times can lead to low levels of peptide backbone hydrolysis.

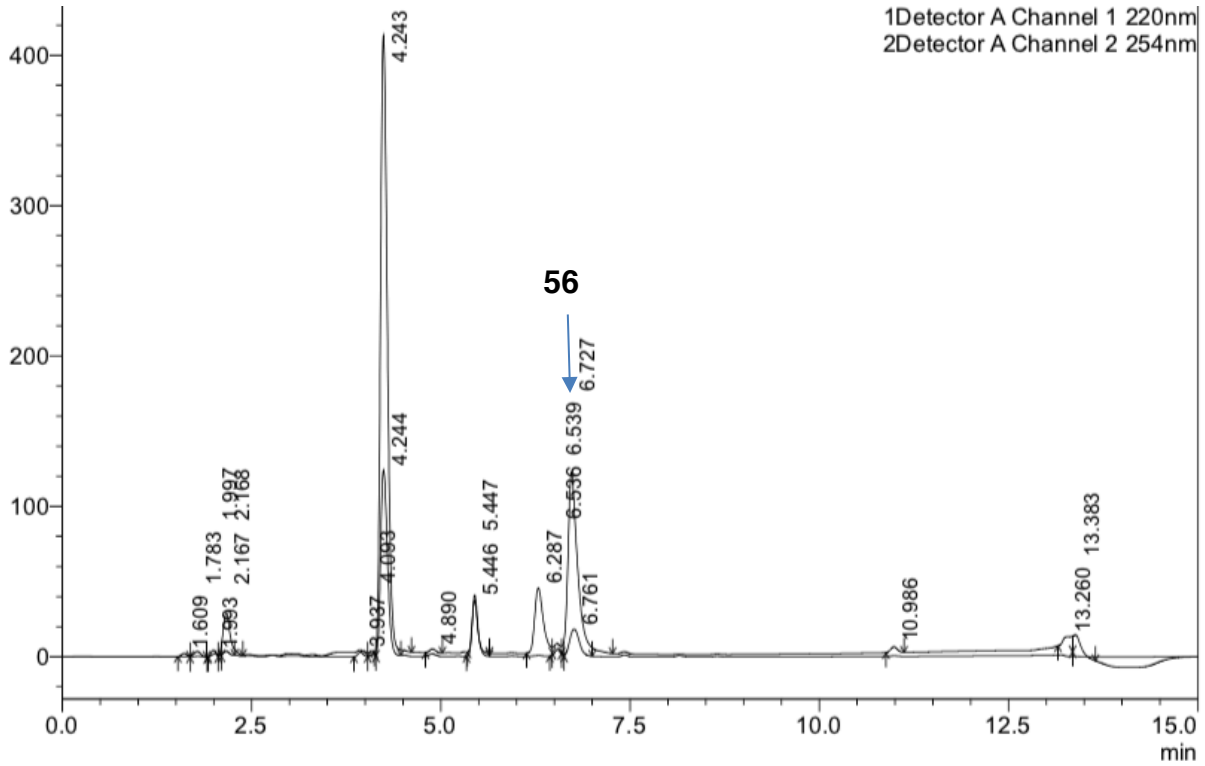


Figure 19 HPLC trace of the crude reaction mixture of exenatide•OAc and **54**

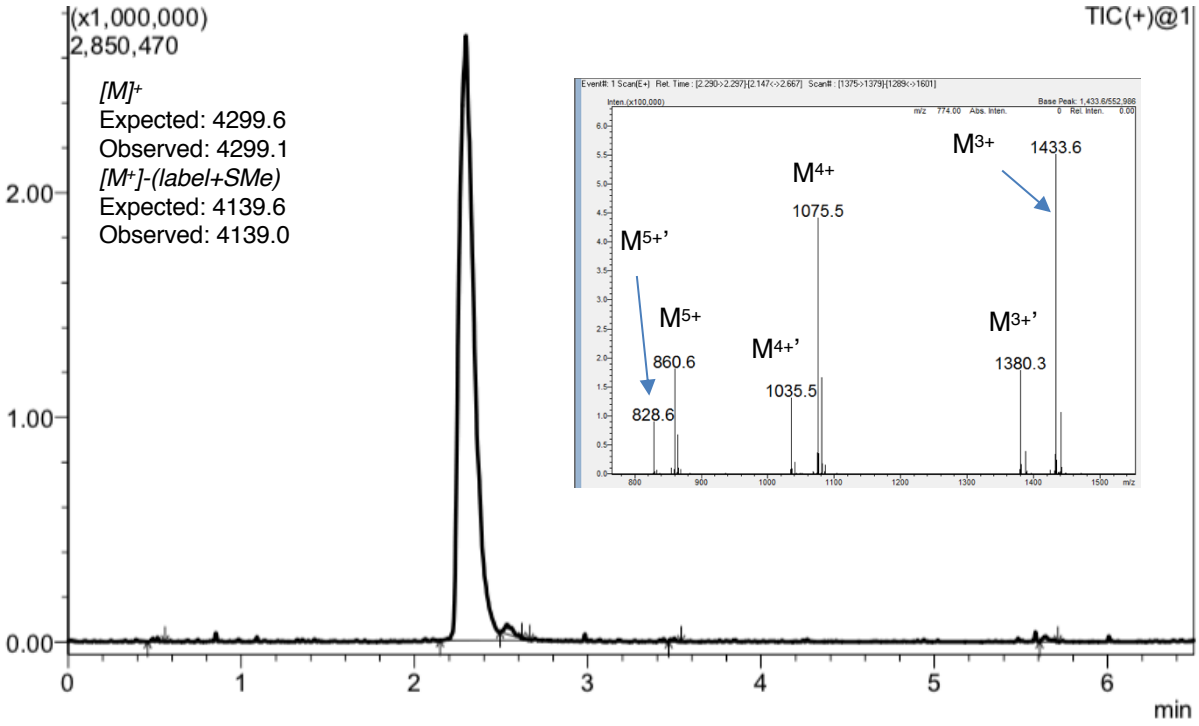
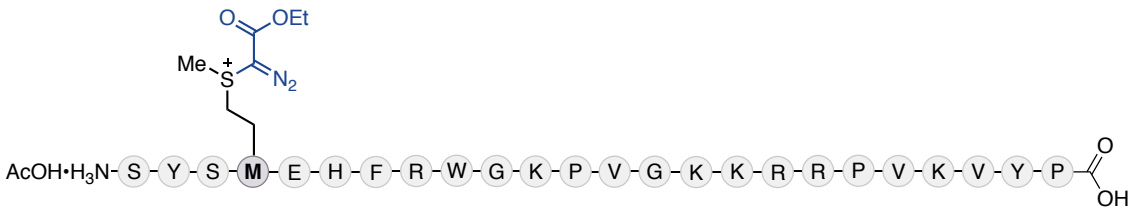


Figure 20 Mass trace of purified exenatide conjugate **56**, with mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with **'**) giving: *M+label-(label+SMe) = M+113-160*.

Tetracosactide functionalization with **54** *in aqueous conditions to form* **57**



The labelling of tetracosactide was performed at room temperature using General Procedure D using **54** (100 mM in H_2O) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) for the extraction. The resulting solution was then analysed directly via LC/MS (C18 column, 50 × 2.1 mm, 2.6 µm, Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min) and was observed to have proceeded in >95% conversion.

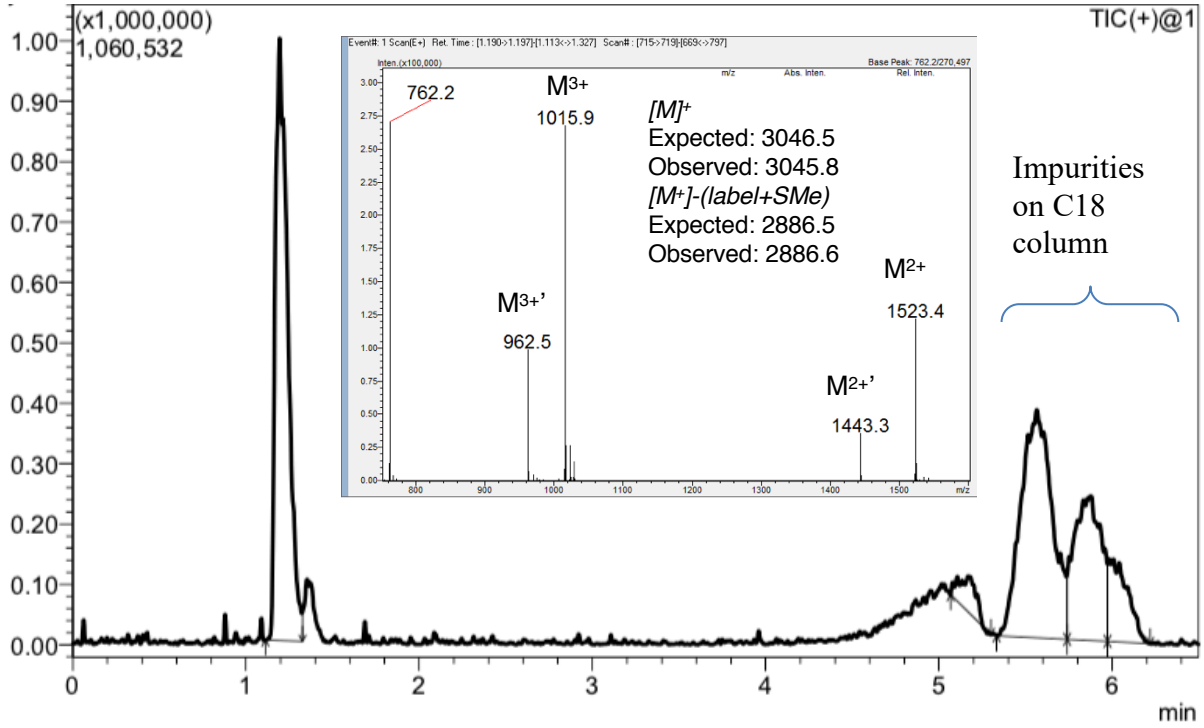


Figure 21 Mass trace of tetracosactide conjugate **57**, with mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with **'**) giving: *M+label-(label+SMe) = M+113-160*.

Large Scale Synthesis of **57**

A 2 mL vial equipped with a magnetic stirrer was charged with Tetracosactide•OAc (~1.8 mg, ~0.6 µmol). Thiourea (100 mM in H_2O, 160 µL), H_2O (120 µL) and TEMPO (67 mM in H_2O, 120 µL) were added and the vial stirred vigorously before **54** (100 mM in H_2O, 400 µL) was added and the resulting solution stirred for 5 minutes at room temperature. The reaction mixture was extracted with 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) and the remaining organic volatiles were then removed from the aqueous layer in vacuo. Purification via reverse-phase HPLC* (C18 column, 150 × 10 mm, 10µm, Gradient: 5-55% B over 9 min then 55-95% B over 1 min, hold for 1.5 min, then 95-5%B over 0.5 min, hold for 3 min, 5 ml/min), followed by lyophilization yielded **57** as a sticky, off-white solid that was immediately redissolved in H_2O (500 µL) and stored at –20 °C. Concentration analysis by A_{280} indicated a 1.0 mg/mL solution ~56% isolated yield).

*Note: Purification of **57** should be carried out expediently as prolonged purification times can lead to low levels of peptide backbone hydrolysis.

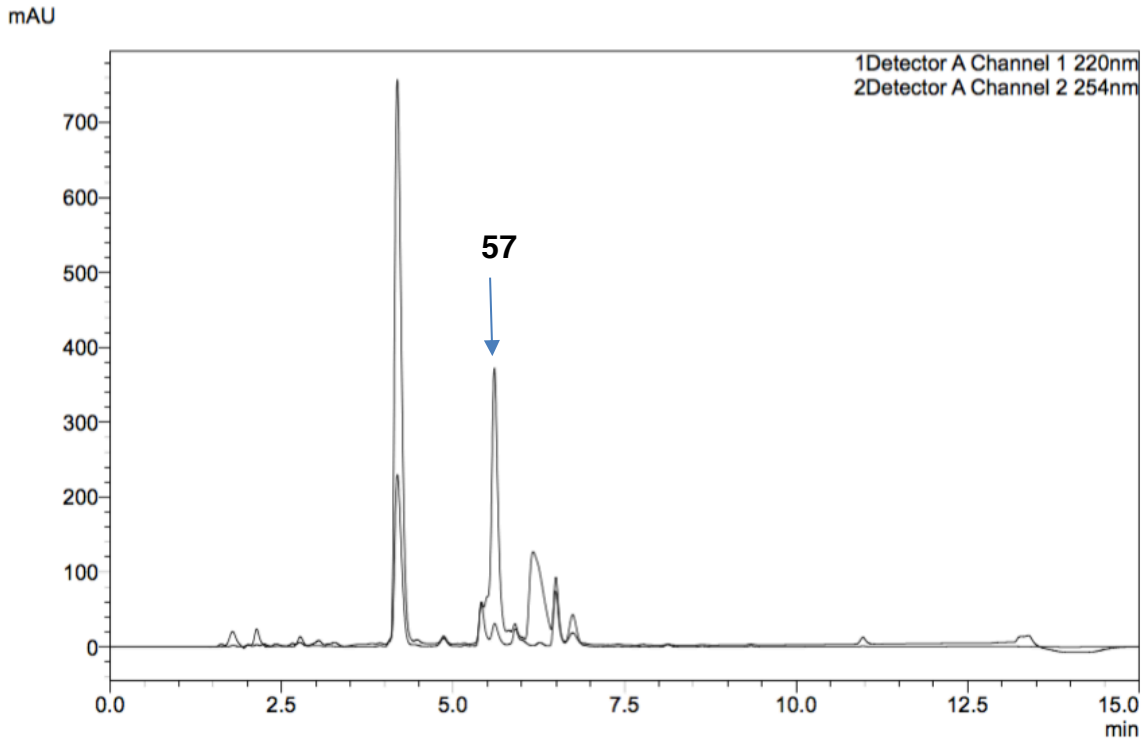


Figure 22 HPLC trace of the crude reaction mixture of tetracosactide•OAc and **54**

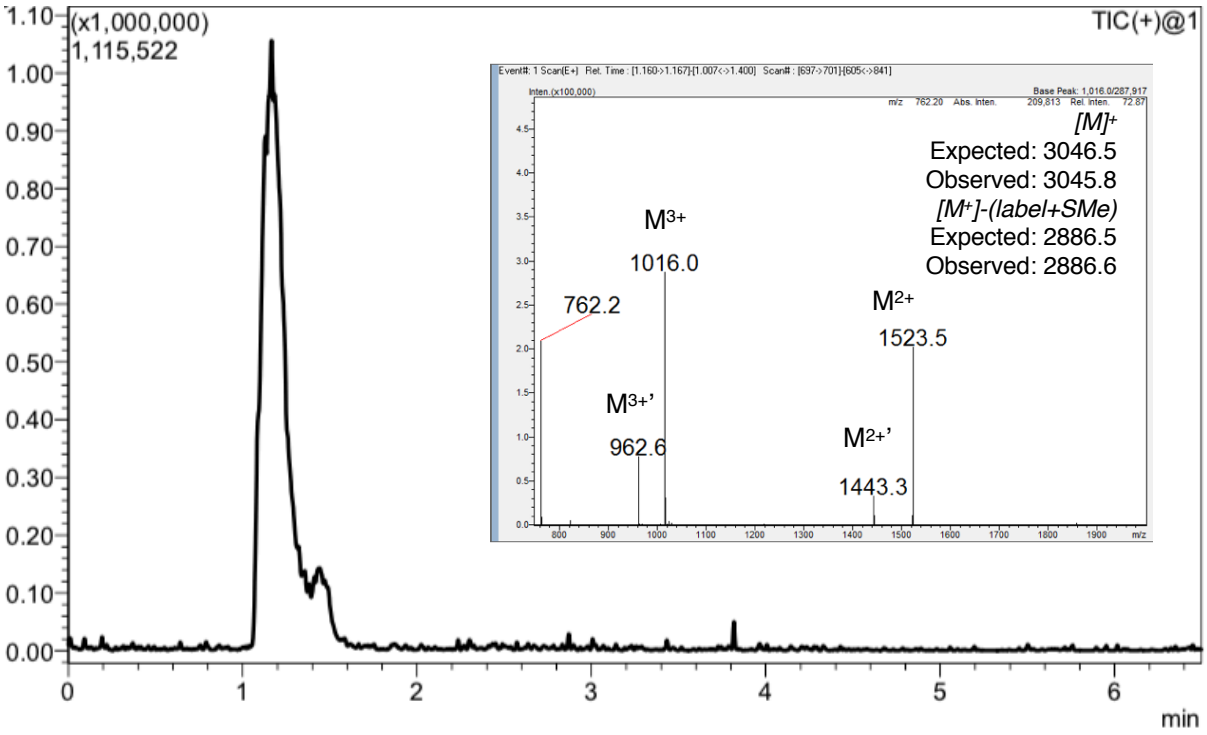


Figure 23 Mass trace of purified tetracosactide conjugate **57**, with mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with ') giving: *M+label-(label+SMe) = M+113-160.*

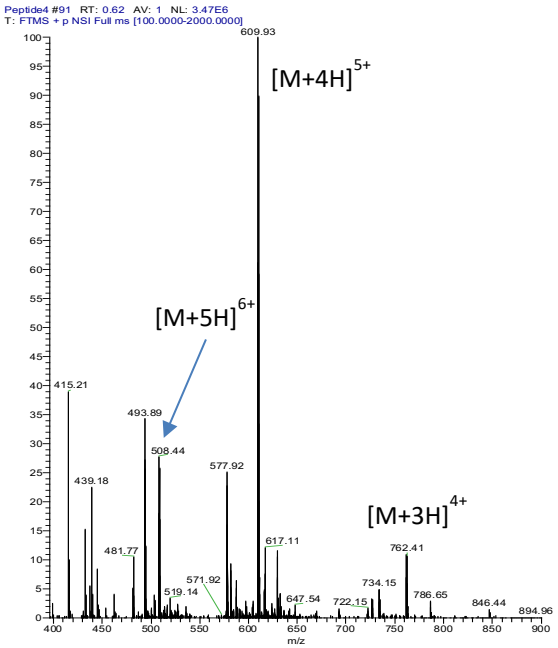
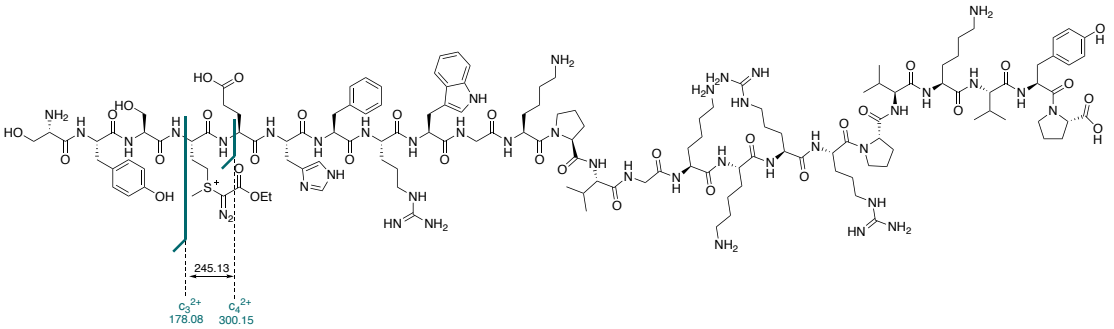
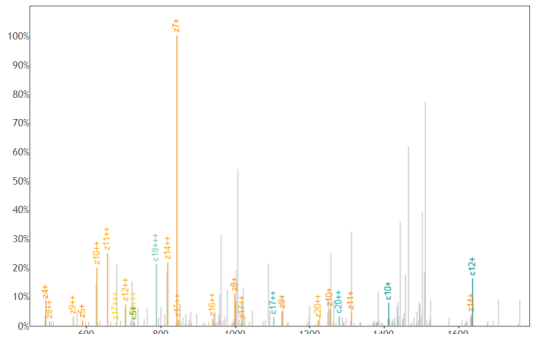


Figure 24 MS Scan of tetracosactide conjugate **57**

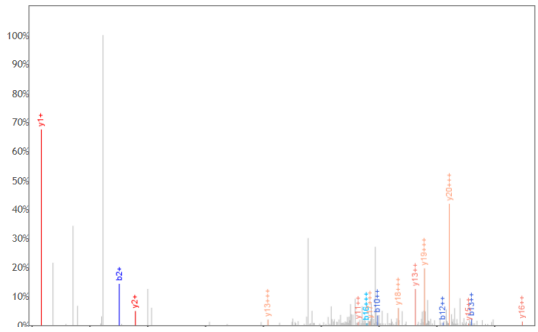


Sequence: SYS**M**EHFRWGKPVGKKRRPVKVYP

b+	b2+	b3+	c+	c2+	c3+	#	Seq	#	y+	y2+	y3+	z+	z2+	z3+
88.0393	44.5233	30.0180	105.0659	53.0366	35.6935	1	S	24						
251.1026	126.0550	84.3724	268.1292	134.5682	90.0479	2	Y	23	2958.6508	1479.8290	986.8884	2942.6321	1471.8197	981.5489
338.1347	169.5710	113.3831	355.1612	178.0842	119.0586	3	S	22	2795.5875	1398.2974	932.5340	2779.5687	1390.2880	927.1944
582.2700	291.6387	194.7615	599.2966	300.1519	200.4370	4	M	21	2708.5554	1354.7814	903.5233	2692.5367	1346.7720	898.1838
711.3126	356.1600	237.7757	728.3392	364.6732	243.4512	5	E	20	2464.4200	1232.7137	822.1449	2448.4013	1224.7043	816.8053
848.3716	424.6894	283.4620	865.3981	433.2027	289.1376	6	H	19	2335.3775	1168.1924	779.1307	2319.3587	1160.1830	773.7911
995.4400	498.2236	332.4848	1012.4665	506.7369	338.1604	7	F	18	2198.3185	1099.6629	733.4444	2182.2998	1091.6535	728.1048
1151.5411	576.2742	384.5185	1168.5676	584.7875	390.1941	8	R	17	2051.2501	1026.1287	684.4216	2035.2314	1018.1193	679.0820
1337.6204	669.3138	446.5450	1354.6469	677.8271	452.2205	9	W	16	1895.1490	948.0781	632.3879	1879.1303	940.0688	627.0483
1394.6419	697.8246	465.5521	1411.6684	706.3378	471.2277	10	G	15	1709.0697	855.0385	570.3614	1693.0510	847.0291	565.0218
1522.7368	761.8720	508.2505	1539.7634	770.3853	513.9260	11	K	14	1652.0482	826.5278	551.3543	1636.0295	818.5184	546.0147
1619.7896	810.3984	540.6014	1636.8161	818.9117	546.2769	12	P	13	1523.9533	762.4803	508.6559	1507.9346	754.4709	503.3164
1718.8580	859.9326	573.6242	1735.8845	868.4459	579.2997	13	V	12	1426.9005	713.9539	476.3050	1410.8818	705.9445	470.9654
1775.8795	888.4434	592.6313	1792.9060	896.9566	598.3069	14	G	11	1327.8321	664.4197	443.2822	1311.8134	656.4103	437.9426
1903.9744	952.4908	635.3297	1921.0010	961.0041	641.0052	15	K	10	1270.8106	635.9090	424.2751	1254.7919	627.8996	418.9355
2032.0694	1016.5383	678.0280	2049.0959	1025.0516	683.7035	16	K	9	1142.7157	571.8615	381.5767	1126.6969	563.8521	376.2372
2188.1705	1094.5889	730.0617	2205.1970	1103.1022	735.7372	17	R	8	1014.6207	507.8140	338.8784	998.6020	499.8046	333.5388
2344.2716	1172.6394	782.0954	2361.2982	1181.1527	787.7709	18	R	7	858.5196	429.7634	286.8447	842.5009	421.7541	281.5051
2441.3244	1221.1658	814.4463	2458.3509	1229.6791	820.1218	19	P	6	702.4185	351.7129	234.8110	686.3998	343.7035	229.4714
2540.3928	1270.7000	847.4691	2557.4193	1279.2133	853.1446	20	V	5	605.3657	303.1865	202.4601	589.3470	295.1771	197.1205
2668.4877	1334.7475	890.1674	2685.5143	1343.2608	895.8429	21	K	4	506.2973	253.6523	169.4373	490.2786	245.6429	164.0977
2767.5562	1384.2817	923.1902	2784.5827	1392.7950	928.8658	22	V	3	378.2023	189.6048	126.7390	362.1836	181.5954	121.3994
2930.6195	1465.8134	977.5447	2947.6460	1474.3267	983.2202	23	Y	2	279.1339	140.0706	93.7162	263.1152	132.0612	88.3766
						24	P	1	116.0706	58.5389	39.3617	100.0519	50.5296	34.0221



ETD fragmentation spectrum c and z ions



HCD fragmentation spectrum b and y ions

Figure 25 MS/MS spectra of tetracosactide conjugate **57** with the difference between c_3^{2+} and c_4^{2+} ions of 245.13 Da confirming the presence of the modification on the methionine residue.

CD Spectrum of tetracosactide conjugate **57**

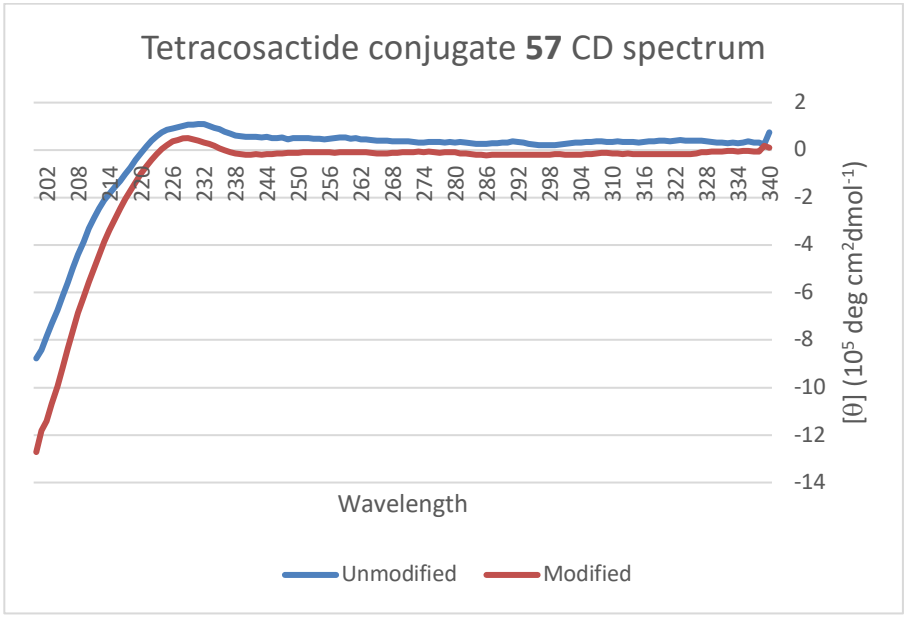
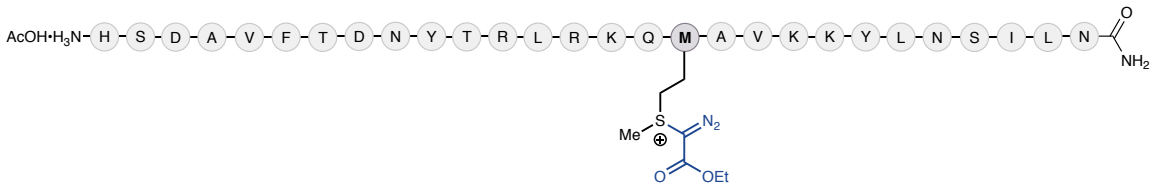


Figure 26 CD spectra of **57** (50 µM) in 5 mM pH 7.4 phosphate buffer, and tetracosactide (25 µM) in 5 mM pH 7.4 phosphate buffer.

Aviptadil functionalization with **54** *in aqueous conditions to form* **58**



The labelling of aviptadil was performed according to General Procedure D using **54** (100 mM in H_2O) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) for the extraction. The resulting solution was then analysed directly via LC/MS (pentafluorophenyl column, 50 × 2.1 mm, 2.6µm, Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min), and was observed to have proceeded in >95% conversion.

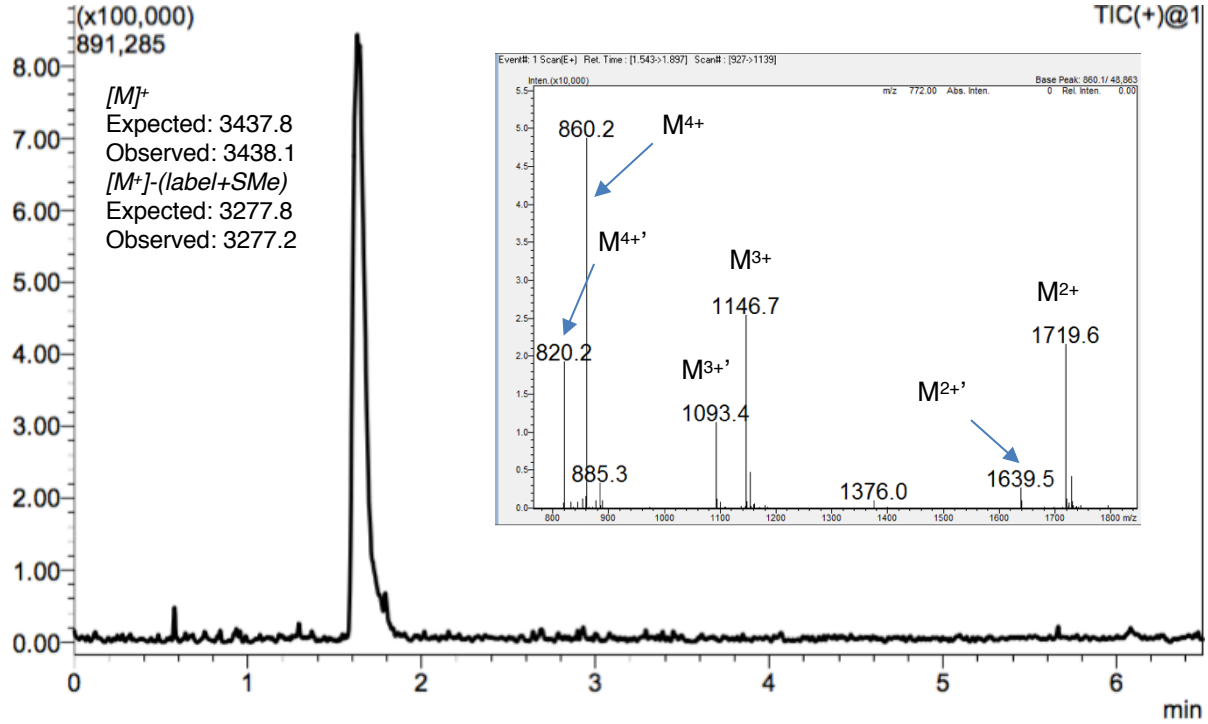


Figure 27 Mass trace of crude reaction mixture of aviptadil with **54**, with mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with ') giving: *M+label-(label+SMe) = M+113-160*.

Large Scale Synthesis of **58**

A 2 mL vial equipped with a magnetic stir bar was charged with Aviptadil•OAc (~1.4 mg, ~0.4 µmol). Thiourea (100 mM in H_2O, 160 µL), H_2O (120 µL) and TEMPO (67 mM in H_2O, 120 µL) were added and the vial was stirred vigorously before **54** (100 mM in H_2O 400 µL) was added. The reaction mixture was stirred for 5 minutes at room temperature before being extracted with 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL). The remaining organic volatiles were then removed in vacuo. Purification via reverse-phase HPLC* (C18 column,150 × 10 mm, 10µm, Gradient: 5-55% B over 9 min then 55-95% B over 1 min, hold for 1.5 min, then 95-5%B over 0.5 min, hold for 3 min, 5 ml/min), followed by lyophilization yielded **58** as a sticky, off-white solid that was immediately redissolved in H_2O (500 µL) and stored at –20 °C. Concentration analysis by A_{280} indicated a 1.3 mg/mL solution (>90% isolated yield).

***Note**: Purification of **58** should be carried out expediently as prolonged purification times can lead to low levels of peptide backbone hydrolysis.

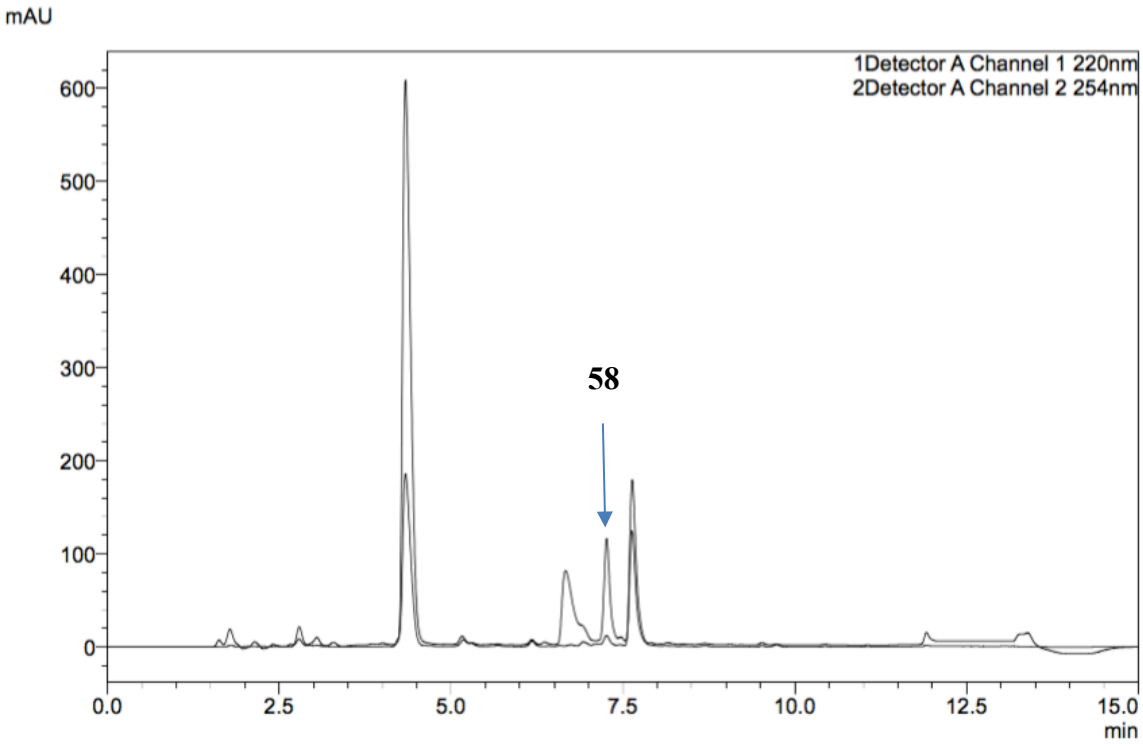


Figure 28 HPLC trace of the crude reaction mixture of aviptadil•OAc and **54**.

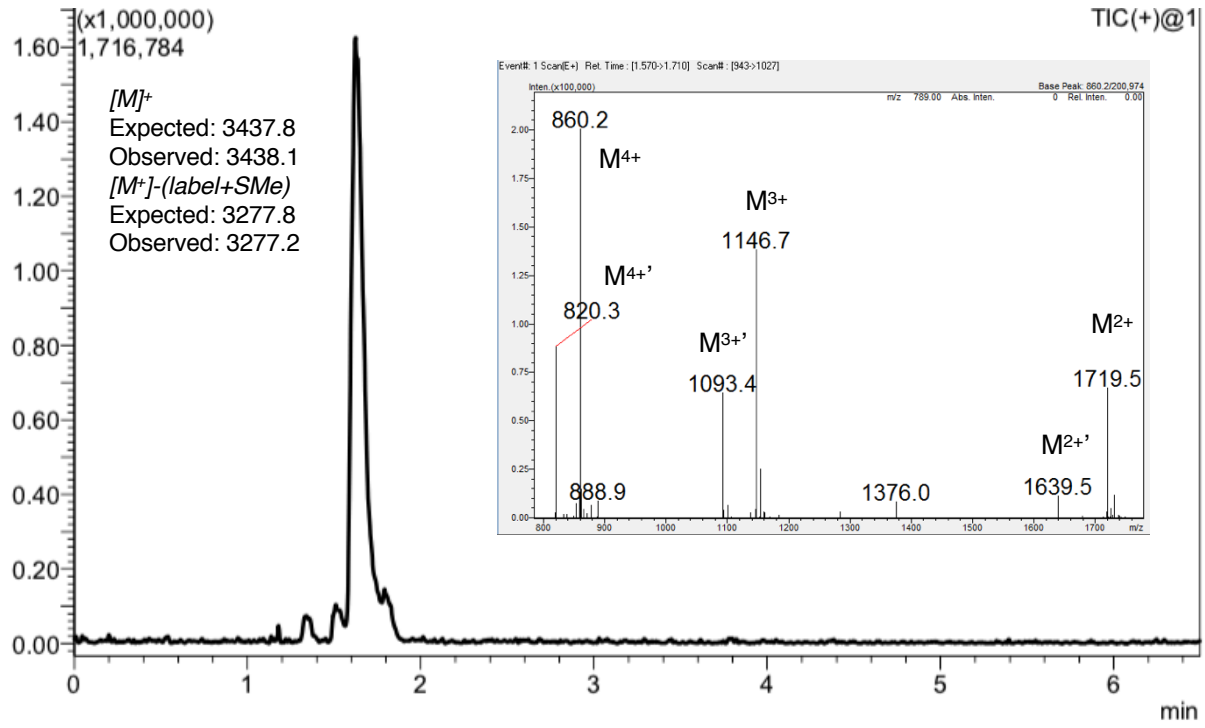
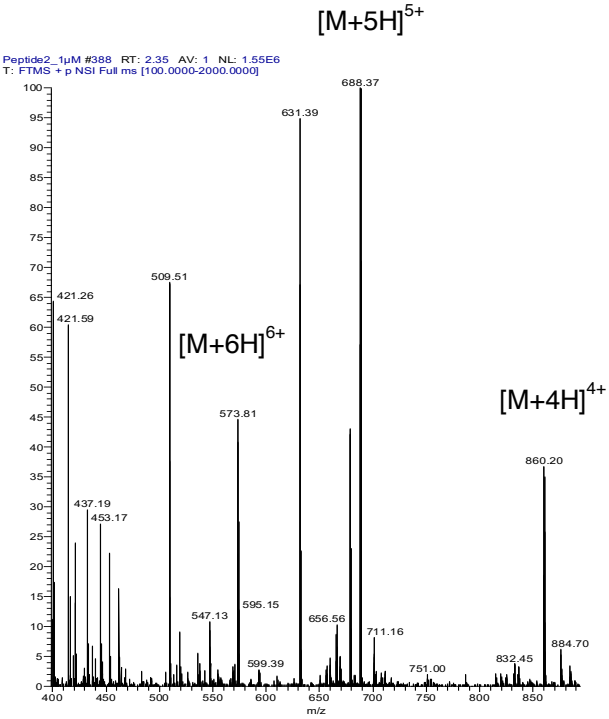
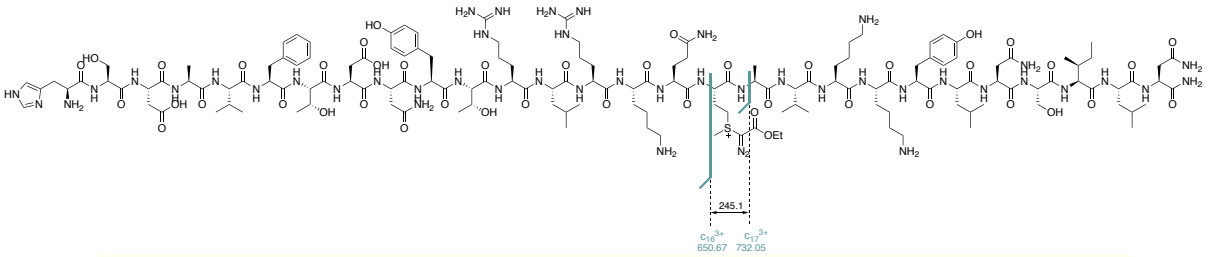


Figure 29 Mass trace of purified aviptadil conjugate **58**, with mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with **'**) giving: *M+label-(label+SMe) = M+113-160*.

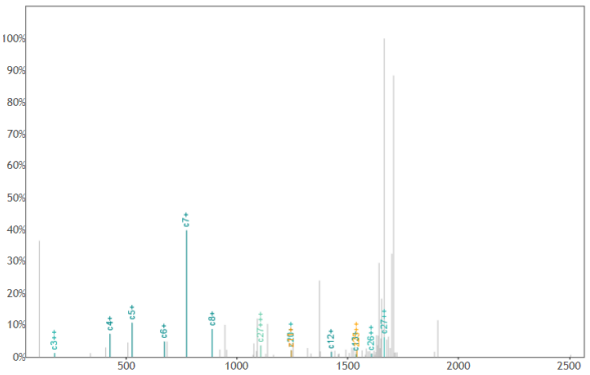


MS scan 1µM **5** in 50% ACN, 0.1%

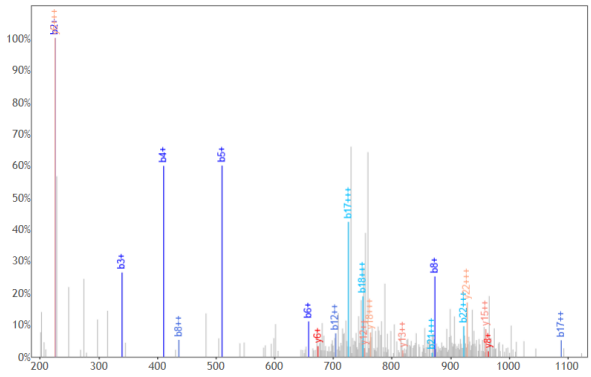
Figure 30 MS of aviptadil conjugate **58**



b+	b2+	b3+	c+	c2+	c3+	#	Seq	#	y+	y2+	y3+	z+	z2+	z3+
138.0662	69.5367	46.6936	155.0927	78.0500	52.3691	1	H	28						
225.0982	113.0527	75.7043	242.1248	121.5660	81.3798	2	S	27	3301.7834	1651.3953	1101.2660	3285.7647	1643.3860	1095.9264
340.1252	170.5662	114.0466	357.1517	179.0795	119.7221	3	D	26	3214.7514	1607.8793	1072.2553	3198.7326	1599.8700	1066.9157
411.1623	206.0848	137.7256	428.1888	214.5980	143.4011	4	A	25	3099.7244	1550.3658	1033.9130	3083.7057	1542.3565	1028.5734
510.2307	255.6190	170.7484	527.2572	264.1323	176.4239	5	V	24	3028.6873	1514.8473	1010.2339	3012.6686	1506.8379	1004.8944
657.2991	329.1532	219.7712	674.3257	337.6665	225.4467	6	F	23	2929.6189	1465.3131	977.2111	2913.6002	1457.3037	971.8716
758.3468	379.6770	253.4538	775.3733	388.1903	259.1293	7	T	22	2782.5505	1391.7789	928.1883	2766.5317	1383.7695	922.8488
873.3737	437.1905	291.7961	890.4003	445.7038	297.4716	8	D	21	2681.5028	1341.2550	894.5058	2665.4841	1333.2457	889.1662
987.4167	494.2120	329.8104	1004.4432	502.7252	335.4859	9	N	20	2566.4758	1283.7416	856.1635	2550.4571	1275.7322	850.8239
1150.4800	575.7436	384.1648	1167.5065	584.2569	389.8404	10	Y	19	2452.4329	1226.7201	818.1492	2436.4142	1218.7107	812.8096
1251.5277	626.2675	417.8474	1268.5542	634.7807	423.5229	11	T	18	2289.3696	1145.1884	763.7947	2273.3509	1137.1791	758.4551
1407.6288	704.3180	469.8811	1424.6553	712.8313	475.5566	12	R	17	2188.3219	1094.6646	730.1122	2172.3032	1086.6552	724.7726
1520.7128	760.8601	507.5758	1537.7394	769.3733	513.2513	13	L	16	2032.2208	1016.6140	678.0785	2016.2021	1008.6047	672.7389
1676.8139	838.9106	559.6095	1693.8405	847.4239	565.2850	14	R	15	1919.1367	960.0720	640.3838	1903.1180	952.0626	635.0442
1804.9089	902.9581	602.3078	1821.9355	911.4714	607.9833	15	K	14	1763.0356	882.0214	588.3501	1747.0169	874.0121	583.0105
1932.9675	966.9874	644.9940	1949.9940	975.5007	650.6695	16	Q	13	1634.9407	817.9740	545.6517	1618.9219	809.9646	540.3122
2177.1029	1089.0551	726.3725	2194.1294	1097.5683	732.0480	17	M	12	1506.8821	753.9447	502.9655	1490.8634	745.9353	497.6260
2248.1400	1124.5736	750.0515	2265.1665	1133.0869	755.7270	18	A	11	1262.7467	631.8770	421.5871	1246.7280	623.8676	416.2475
2347.2084	1174.1078	783.0743	2364.2349	1182.6211	788.7498	19	V	10	1191.7096	596.3584	397.9080	1175.6909	588.3491	392.5685
2475.3034	1238.1553	825.7726	2492.3299	1246.6686	831.4482	20	K	9	1092.6412	546.8242	364.8852	1076.6224	538.8149	359.5457
2603.3983	1302.2028	868.4710	2620.4249	1310.7161	874.1465	21	K	8	964.5462	482.7767	322.1869	948.5275	474.7674	316.8473
2766.4617	1383.7345	922.8254	2783.4882	1392.2477	928.5009	22	Y	7	836.4512	418.7293	279.4886	820.4325	410.7199	274.1490
2879.5457	1440.2765	960.5201	2896.5723	1448.7898	966.1956	23	L	6	673.3879	337.1976	225.1342	657.3692	329.1882	219.7946
2993.5886	1497.2980	998.5344	3010.6152	1505.8112	1004.2099	24	N	5	560.3039	280.6556	187.4395	544.2851	272.6462	182.0999
3080.6207	1540.8140	1027.5451	3097.6472	1549.3272	1033.2206	25	S	4	446.2609	223.6341	149.4252	430.2422	215.6247	144.0856
3193.7047	1597.3560	1065.2398	3210.7313	1605.8693	1070.9153	26	I	3	359.2289	180.1181	120.4145	343.2102	172.1087	115.0749
3306.7888	1653.8980	1102.9345	3323.8153	1662.4113	1108.6100	27	L	2	246.1448	123.5761	82.7198	230.1261	115.5667	77.3802
						28	N	1	133.0608	67.0340	45.0251	117.0420	59.0247	39.6855



ETD fragmentation spectrum c and z ions



HCD fragmentation spectrum b and y ions

Figure 31 MS/MS spectra of aviptadil conjugate **58** with the difference between c_{16}^{3+} and c_{17}^{3+} ions of 245.1 Da confirming the presence of the modification on the methionine residue.

CD Spectrum of aviptadil conjugate **58**

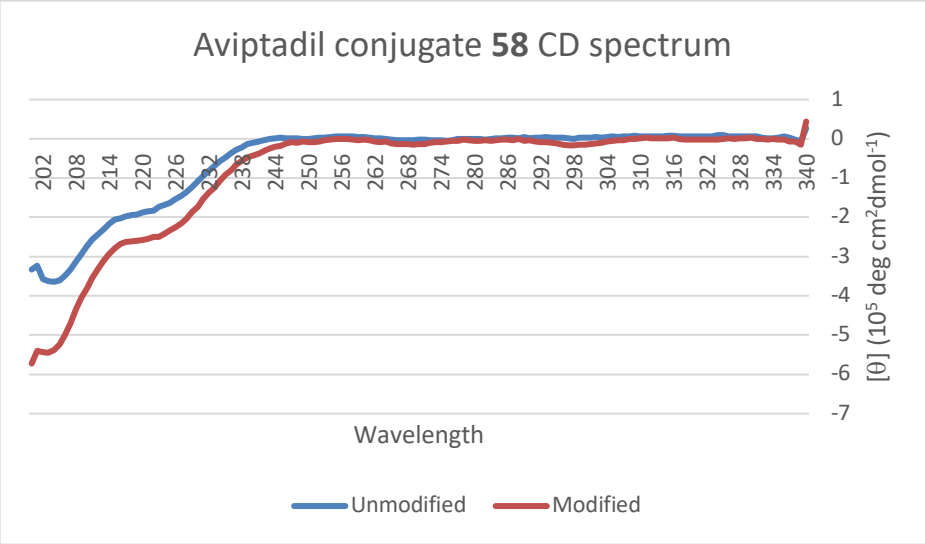
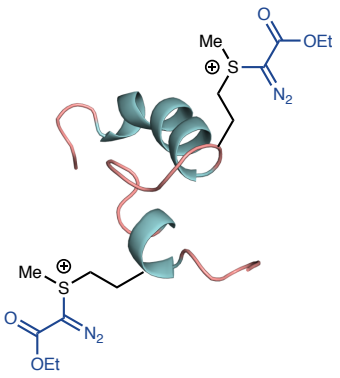


Figure 32 CD spectra of **58** (50 µM) and aviptadil (25 µM), both in a 5 mM pH 7.4 phosphate buffer.

Teriparatide functionalization with **54** *in aqueous conditions to form* **59**



The labelling of teriparatide was performed at room temperature using General Procedure E using **54** (100 mM in H_2O) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) for the extraction. The resulting solution was then analysed directly via LC/MS (C18 column 50 × 2.1 mm, 2.6 µm, Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min) and was observed to have proceeded in >95% conversion.

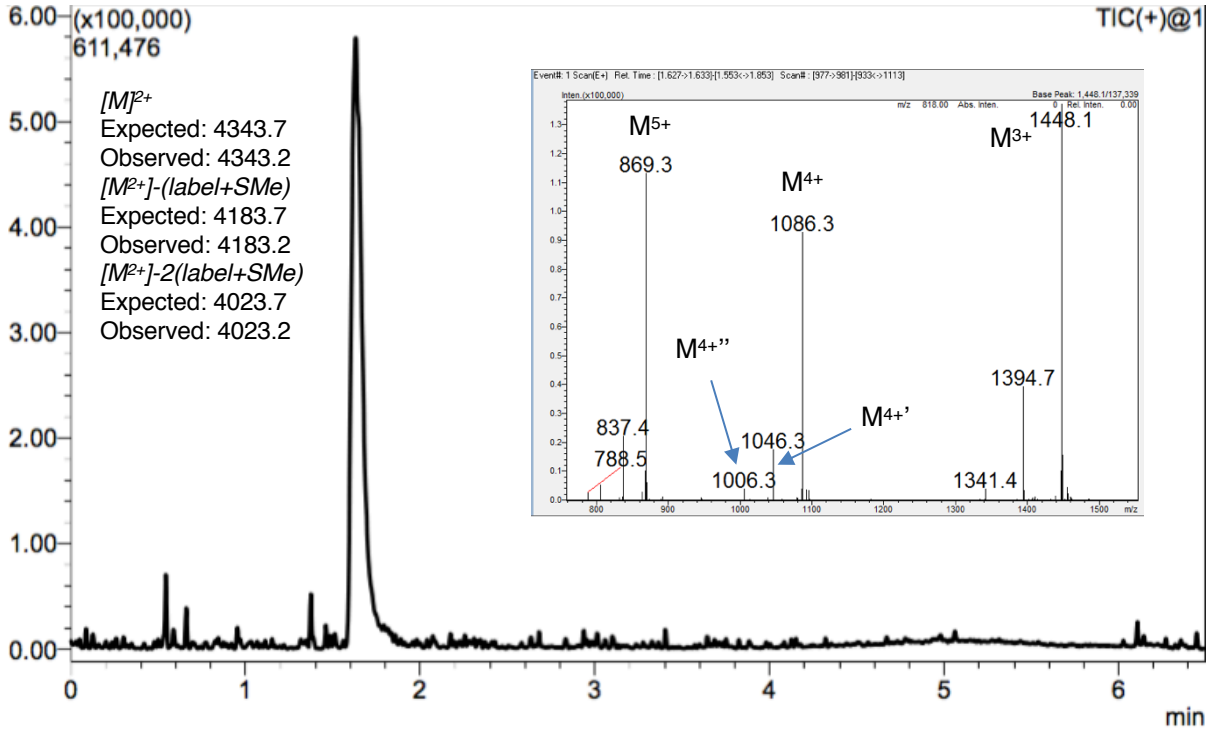


Figure 33 Mass trace of crude reaction mixture of teriparatide with **54**, mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with ‘) giving: *M+2(label)-(label+SMe) = M+2(113)-160.* The *M+2(label)-2(label+SMe) = M+2(113)-2(160)* is indicated by ’’.

Large Scale Synthesis of **59**

A 2 mL vial equipped with a magnetic stirrer was charged with teriparatide•OAc (~1.2 mg, ~0.3 µmol). Thiourea (100 mM in H_2O, 160 µL), formic acid (0.1 M in H_2O, 40 µL), H_2O (80 µL) and TEMPO (67 mM in H_2O, 120 µL) were added and the vial was stirred vigorously before **54** (100 mM in H_2O, 400 µL) was added. The reaction mixture was stirred for 5 minutes at room temperature, then extracted with 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL). The remaining organic volatiles were then removed in vacuo. Purification via reverse-phase HPLC* (C18 column, 150 × 10 mm, 10 µm, Gradient: 5-55% B over 9 min then 55-95% B over 1 min, hold for 1.5 min, then 95-5% B over 0.5 min, hold for 3 min, 5 mL/min), followed by lyophilization yielded **59** as a sticky, off-white solid that was immediately redissolved in H_2O (500 µL) and stored at –20 °C. Concentration analysis by A_{280} indicated a 1.0 mg/mL solution ~92% isolated yield).

***Note**: Purification of **59** should be carried out expediently as prolonged purification times can lead to low levels of peptide backbone hydrolysis.

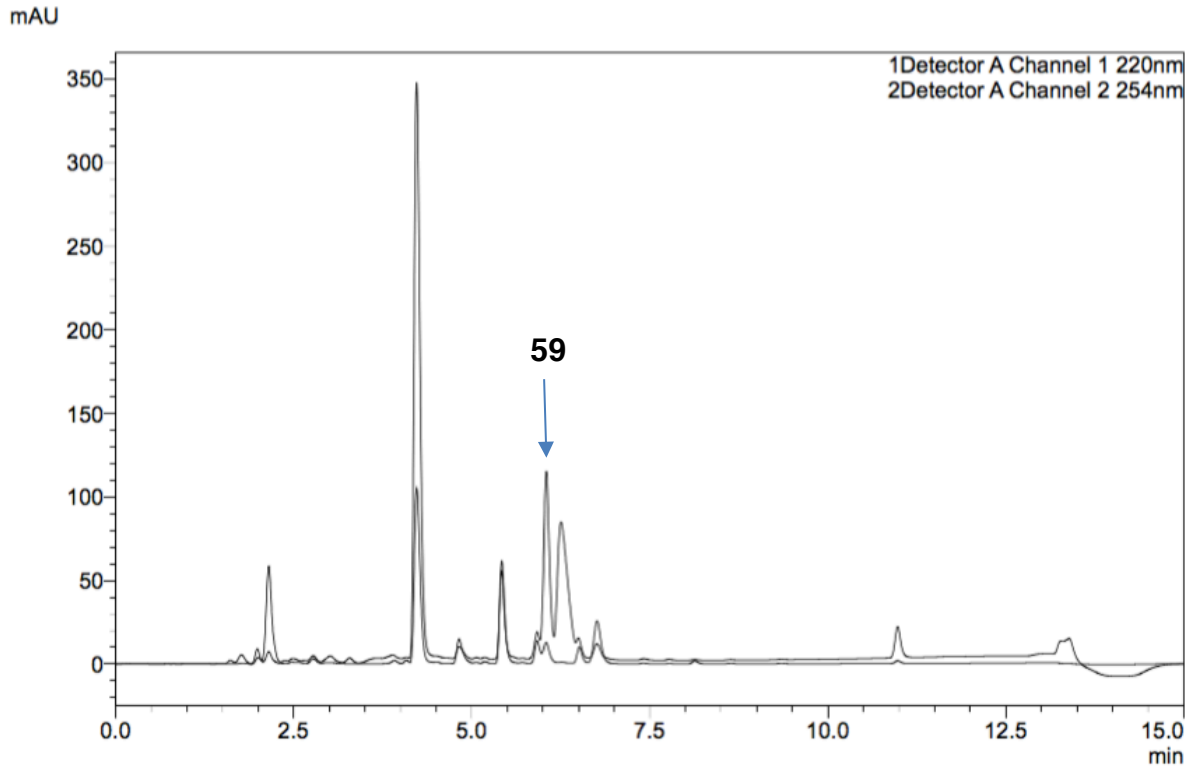


Figure 34 HPLC trace of the crude reaction mixture of teriparatide•OAc and **54**

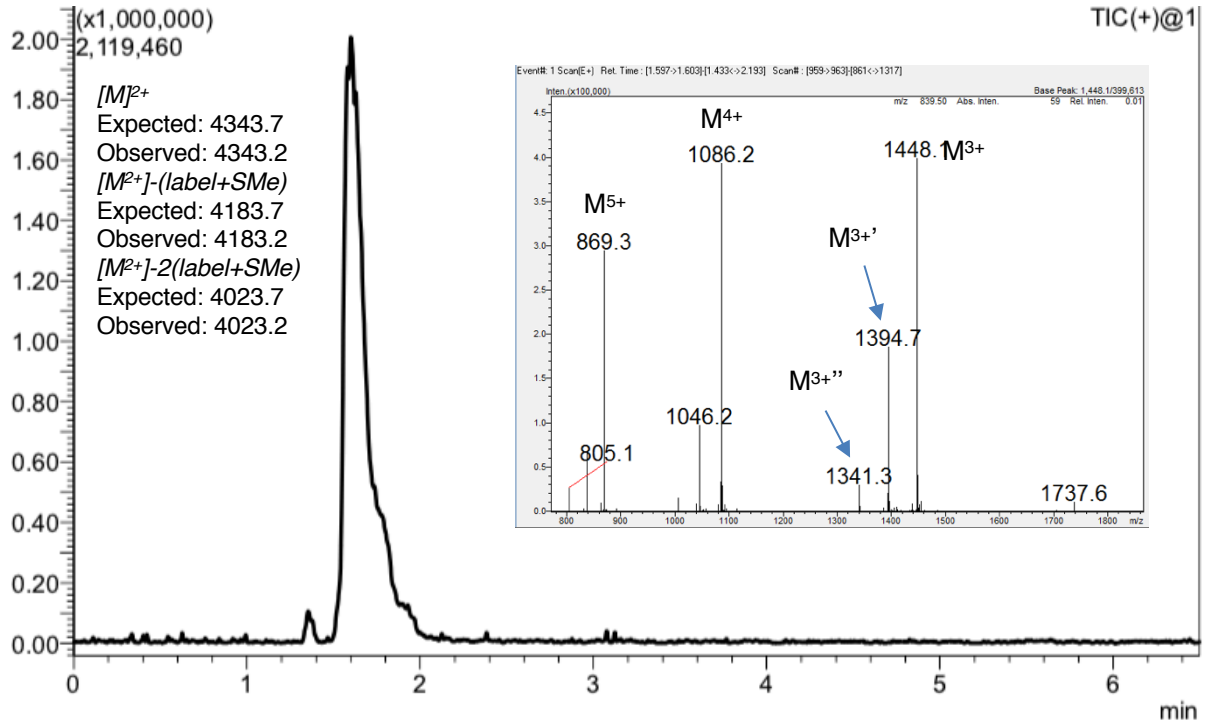
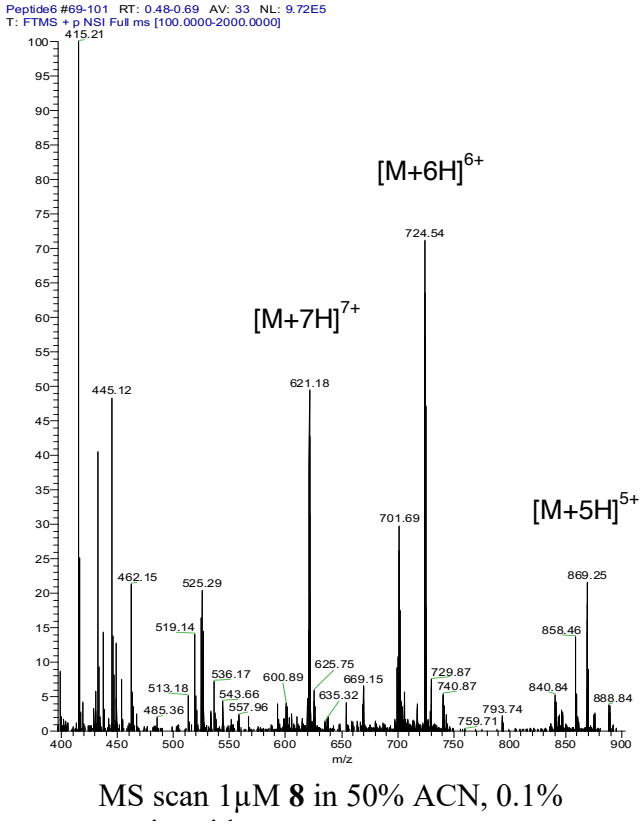
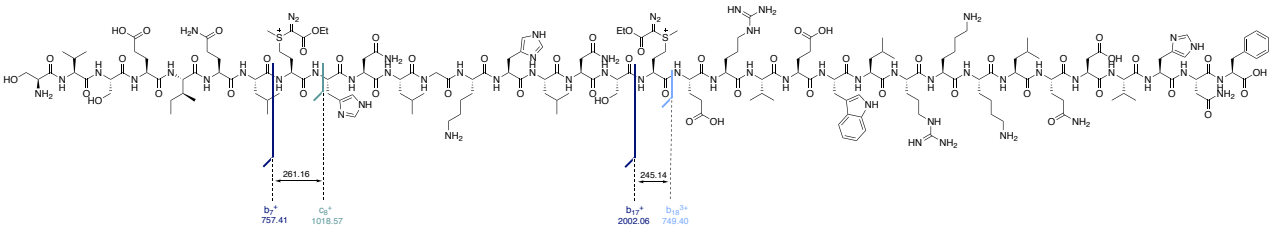


Figure 35 Mass trace of purified teriparatide conjugate **59**, mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with ‘) giving: *M+2(label)-(label+SMe) = M+2(113)-160.* The *M+2(label)-2(label+SMe) = M+2(113)-2(160)* is indicated by ’’.



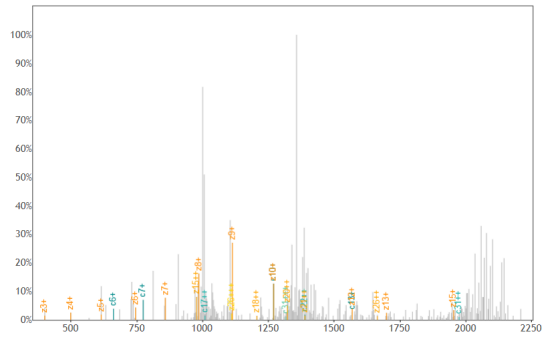
MS scan 1µM **8** in 50% ACN, 0.1%

Figure 36 MS of teriparatide conjugate **59**

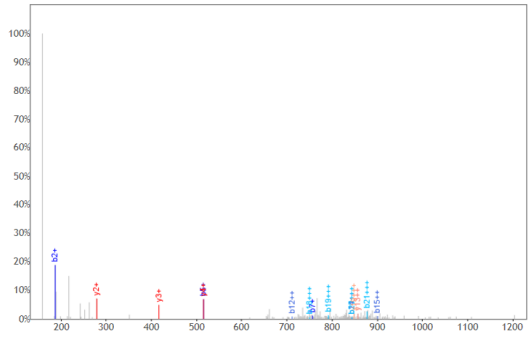


Sequence: SVSEIQL**M**HNLGKHLNS**M**ERVEWLRKKLEDVHNF

b+	b2+	b3+	c+	c2+	c3+	#	Seq	#	y+	y2+	y3+	z+	z2+	z3+
88.0393	44.5233	30.0180	105.0659	53.0366	35.6935	1	S	34						
187.1077	94.0575	63.0408	204.1343	102.5708	68.7163	2	V	33	4256.2800	2128.6436	1419.4315	4240.2613	2120.6343	1414.0919
274.1397	137.5735	92.0514	291.1663	146.0868	97.7269	3	S	32	4157.2116	2079.1094	1386.4087	4141.1929	2071.1001	1381.0691
403.1823	202.0948	135.0656	420.2089	210.6081	140.7411	4	E	31	4070.1796	2035.5934	1357.3980	4054.1608	2027.5841	1352.0585
516.2664	258.6368	172.7603	533.2930	267.1501	178.4358	5	I	30	3941.1370	1971.0721	1314.3838	3925.1182	1963.0628	1309.0443
644.3250	322.6661	215.4465	661.3515	331.1794	221.1220	6	Q	29	3828.0529	1914.5301	1276.6892	3812.0342	1906.5207	1271.3496
757.4090	379.2082	253.1412	774.4356	387.7214	258.8167	7	L	28	3699.9943	1850.5008	1234.0030	3683.9756	1842.4914	1228.6634
1001.5444	501.2759	334.5197	1018.5710	509.7891	340.1952	8	M	27	3586.9103	1793.9588	1196.3083	3570.8915	1785.9494	1190.9687
1138.6033	569.8053	380.2060	1155.6299	578.3186	385.8815	9	H	26	3342.7749	1671.8911	1114.9298	3326.7562	1663.8817	1109.5902
1252.6463	626.8268	418.2203	1269.6728	635.3400	423.8958	10	N	25	3205.7160	1603.3616	1069.2435	3189.6972	1595.3523	1063.9039
1365.7303	683.3688	455.9150	1382.7569	691.8821	461.5905	11	L	24	3091.6730	1546.3402	1031.2292	3075.6543	1538.3308	1025.8896
1422.7518	711.8795	474.9221	1439.7783	720.3928	480.5976	12	G	23	2978.5890	1489.7981	993.5345	2962.5703	1481.7888	988.1949
1550.8468	775.9270	517.6204	1567.8733	784.4403	523.2960	13	K	22	2921.5675	1461.2874	974.5274	2905.5488	1453.2780	969.1878
1687.9057	844.4565	563.3067	1704.9322	852.9697	568.9823	14	H	21	2793.4725	1397.2399	931.8290	2777.4538	1389.2306	926.4895
1800.9897	900.9985	601.0014	1818.0163	909.5118	606.6769	15	L	20	2656.4136	1328.7105	886.1427	2640.3949	1320.7011	880.8032
1915.0327	958.0200	639.0157	1932.0592	966.5332	644.6913	16	N	19	2543.3296	1272.1684	848.4480	2527.3109	1264.1591	843.1085
2002.0647	1001.5360	668.0264	2019.0912	1010.0493	673.7019	17	S	18	2429.2866	1215.1470	810.4337	2413.2679	1207.1376	805.0942
2246.2001	1123.6037	749.4049	2263.2266	1132.1170	755.0804	18	M	17	2342.2546	1171.6309	781.4231	2326.2359	1163.6216	776.0835
2375.2427	1188.1250	792.4191	2392.2692	1196.6382	798.0946	19	E	16	2098.1192	1049.5633	700.0446	2082.1005	1041.5539	694.7050
2531.3438	1266.1755	844.4528	2548.3703	1274.6888	850.1283	20	R	15	1969.0766	985.0420	657.0304	1953.0579	977.0326	651.6908
2630.4122	1315.7097	877.4756	2647.4387	1324.2230	883.1511	21	V	14	1812.9755	906.9914	604.9967	1796.9568	898.9820	599.6571
2759.4548	1380.2310	920.4898	2776.4813	1388.7443	926.1653	22	E	13	1713.9071	857.4572	571.9739	1697.8884	849.4478	566.6343
2945.5341	1473.2707	982.5162	2962.5607	1481.7840	988.1917	23	W	12	1584.8645	792.9359	528.9597	1568.8458	784.9265	523.6201
3058.6182	1529.8127	1020.2109	3075.6447	1538.3260	1025.8864	24	L	11	1398.7852	699.8962	466.9333	1382.7665	691.8869	461.5937
3214.7193	1607.8633	1072.2446	3231.7458	1616.3766	1077.9201	25	R	10	1285.7011	643.3542	429.2386	1269.6824	635.3448	423.8990
3342.8142	1671.9108	1114.9429	3359.8408	1680.4240	1120.6184	26	K	9	1129.6000	565.3037	377.2049	1113.5813	557.2943	371.8653
3470.9092	1735.9582	1157.6413	3487.9358	1744.4715	1163.3168	27	K	8	1001.5051	501.2562	334.5065	985.4863	493.2468	329.1670
3583.9933	1792.5003	1195.3359	3601.0198	1801.0135	1201.0115	28	L	7	873.4101	437.2087	291.8082	857.3914	429.1993	286.4686
3713.0359	1857.0216	1238.3501	3730.0624	1865.5348	1244.0257	29	E	6	760.3260	380.6667	254.1135	744.3073	372.6573	248.7740
3828.0628	1914.5350	1276.6925	3845.0894	1923.0483	1282.3680	30	D	5	631.2835	316.1454	211.0993	615.2647	308.1360	205.7598
3927.1312	1964.0692	1309.7153	3944.1578	1972.5825	1315.3908	31	V	4	516.2565	258.6319	172.7570	500.2378	250.6225	167.4174
4064.1901	2032.5987	1355.4016	4081.2167	2041.1120	1361.0771	32	H	3	417.1881	209.0977	139.7342	401.1694	201.0883	134.3946
4178.2331	2089.6202	1393.4159	4195.2596	2098.1334	1399.0914	33	N	2	280.1292	140.5682	94.0479	264.1105	132.5589	88.7083
						34	F	1	166.0863	83.5468	56.0336	150.0675	75.5374	50.6940



ETD fragmentation spectrum c and z ions



HCD fragmentation spectrum b and y ions

Figure 37 MS/MS spectra of teriparatide conjugate **59** with the difference between b_7^+ and c_8^+ ions of 261.2 Da and difference between b_{17}^+ and b_{18}^{3+} ions of 245.1 Da confirming the presence of the modification on the methionine residues.

CD Spectrum of teriparatide conjugate **59**

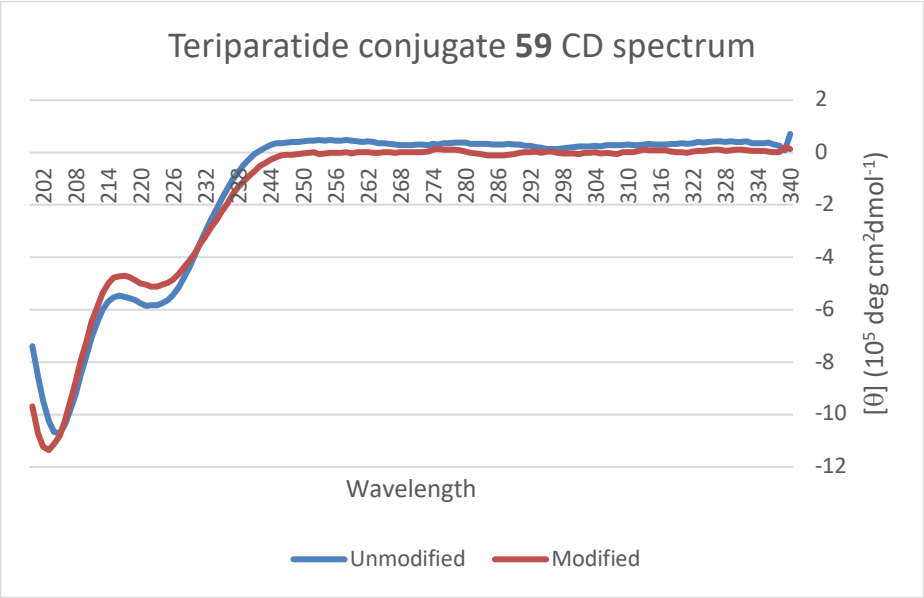


Figure 38 CD spectra of **59** (50 µM) and teriparatide (25 µM), both in 5 mM pH 7.4 phosphate buffer

Large Scale Synthesis of Ubiquitin conjugate **63**

A 2 mL vial equipped with a magnetic stirrer was charged with Ubiquitin (~1.5 mg, ~0.2 µmol). To the vial was added thiourea (100 mM in H_2O, 100 µL), formic acid (0.1M in H_2O, 25 µL) and H_2O (125 µL) and the vial stirred vigorously before **54** (100 mM in H_2O, 250 µL) was added. The reaction mixture was stirred for 5 minutes at room temperature, then extracted twice with 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL). The remaining organic volatiles were then removed from the aqueous layer using a rotary evaporator. **63** was purified by ion exchange chromatography (MonoS 4.6/100 PE column, GE Healthcare) using a NaCl gradient (0-1 M) in 50 mM ammonium acetate pH 4.5 on an ÄKTAexplorer (Pharmacia Biotech), and fractions directly analysed my MS.

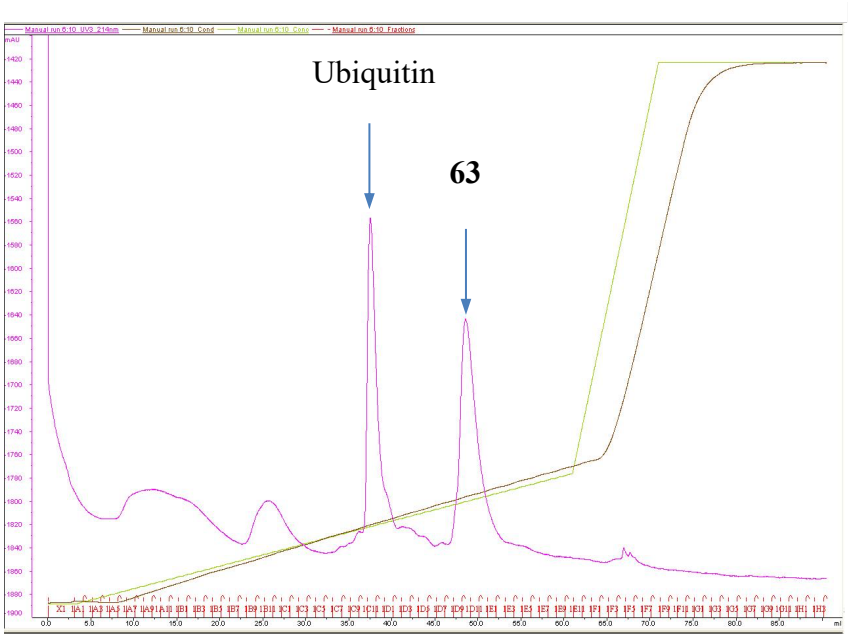


Figure 39 FPLC trace of Ubiquitin conjugate **63**, purified using MonoS 4.6/100 PE column, GE Healthcare. Clear separation of unmodified Ubiquitin and modified conjugate **63** was achieved due to the differing elution times of the charged species **63** and uncharged, unmodified ubiquitin using an ion exchange column. FPLC run with the assistance of Dr N. Hugenin.

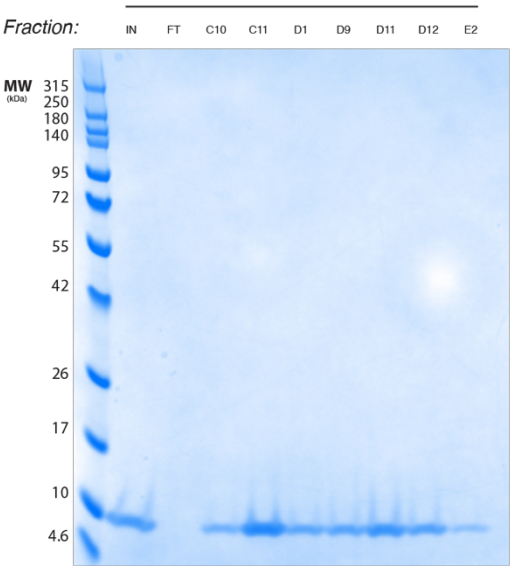
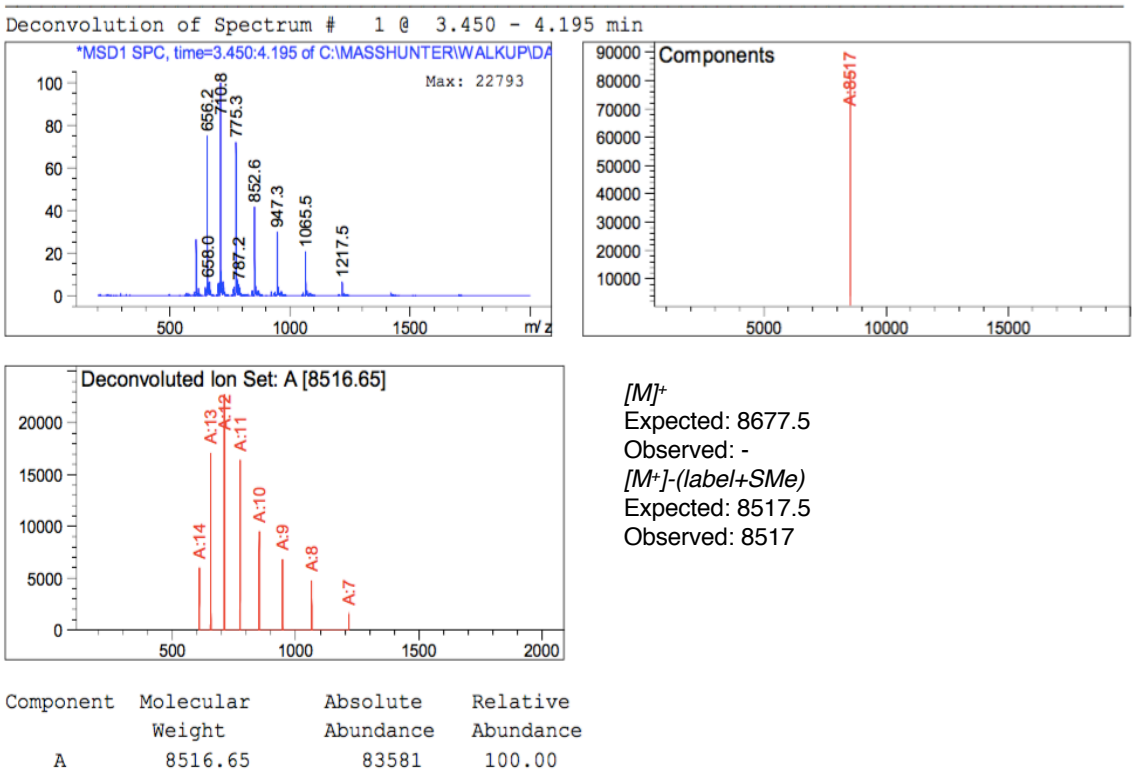


Figure 40 Gel of purified Ubiquitin fractions from FPLC purification run with Bolt 4-12% Bis-Tris gels (Thermo Scientific), run for 30 min at 165V (constant) using MES buffer (Thermo Scientific). The ladder on the gel is ProSieve QuadColor protein marker (4.6 kDa – 300 kDa; Lonza) and the gel was stained using InstantBlue (Generon). Run with the assistance of Dr N. Hugenin.

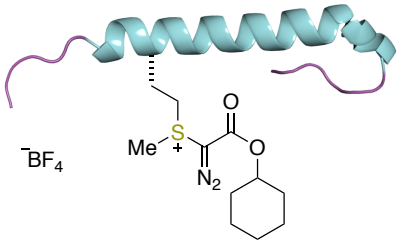


[M]⁺
Expected: 8677.5
Observed: -
[M⁺]-(label+SMe)
Expected: 8517.5
Observed: 8517

Figure 41 Protein LCMS and deconvoluted MS of purified ubiquitin conjugate **63** after FPLC purification. Deconvoluted MS shows the fragment ion of **63**-160 for $[M-(RSMe)]^+$

5.3.2 Substrate scope of hypervalent iodine reagents for polypeptide functionalisation

Exenatide acetate functionalization with **77** *to form* **86**



The labelling of exenatide was performed at room temperature using General Procedure D using **77** (100 mM in H_2O) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) for the extraction. The resulting solution was then analysed directly *via* LC/MS (C18 column, 50 × 2.1 mm, 2.6 µm, Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 mL/min) and was observed to have proceeded in >95% conversion.

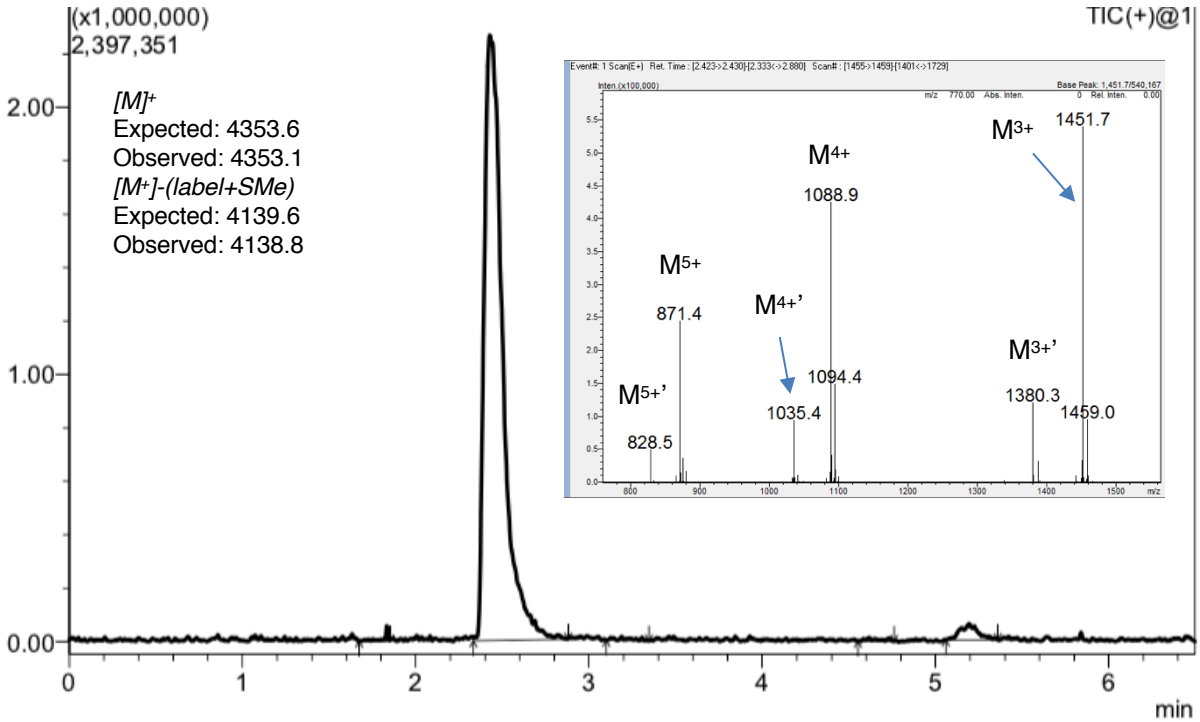


Figure 42 Mass trace of crude reaction mixture of exenatide with **77**, mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with ') giving: *M+label-(label+SMe) = M+167-214*.

Large Scale Synthesis of **86**

A 2 mL vial equipped with a magnetic stirrer was charged with Exenatide•OAc (~1.2 mg, ~0.3 µmol). Thiourea (100 mM in H_2O, 160 µL), formic acid (0.1 M in H_2O, 40 µL), H_2O (80 µL) and TEMPO (67 mM in H_2O, 120 µL) were added and the vial was stirred vigorously before the addition of **77** (100 mM in H_2O, 400 µL). The reaction mixture was stirred for 5 minutes at room temperature before being extracted with 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL). The remaining organic volatiles were then removed in vacuo. Purification via reverse-phase HPLC* (C18 column 150 × 10 mm, 10µm, Gradient: 5-55% B over 9 min then 55-95% B over 1 min, hold for 1.5 min, then 95-5% B over 0.5 min, hold for 3 min, 5 mL/min), followed by lyophilization yielded **86** as a sticky, off-white solid that was immediately redissolved in H_2O (500 µL) and stored at –20 °C. Concentration analysis by A_{280} indicated a 1.5 mg/mL solution ~63% isolated yield).

***Note**: Purification of **86** should be carried out expediently as prolonged purification times can lead to low levels of peptide backbone hydrolysis.

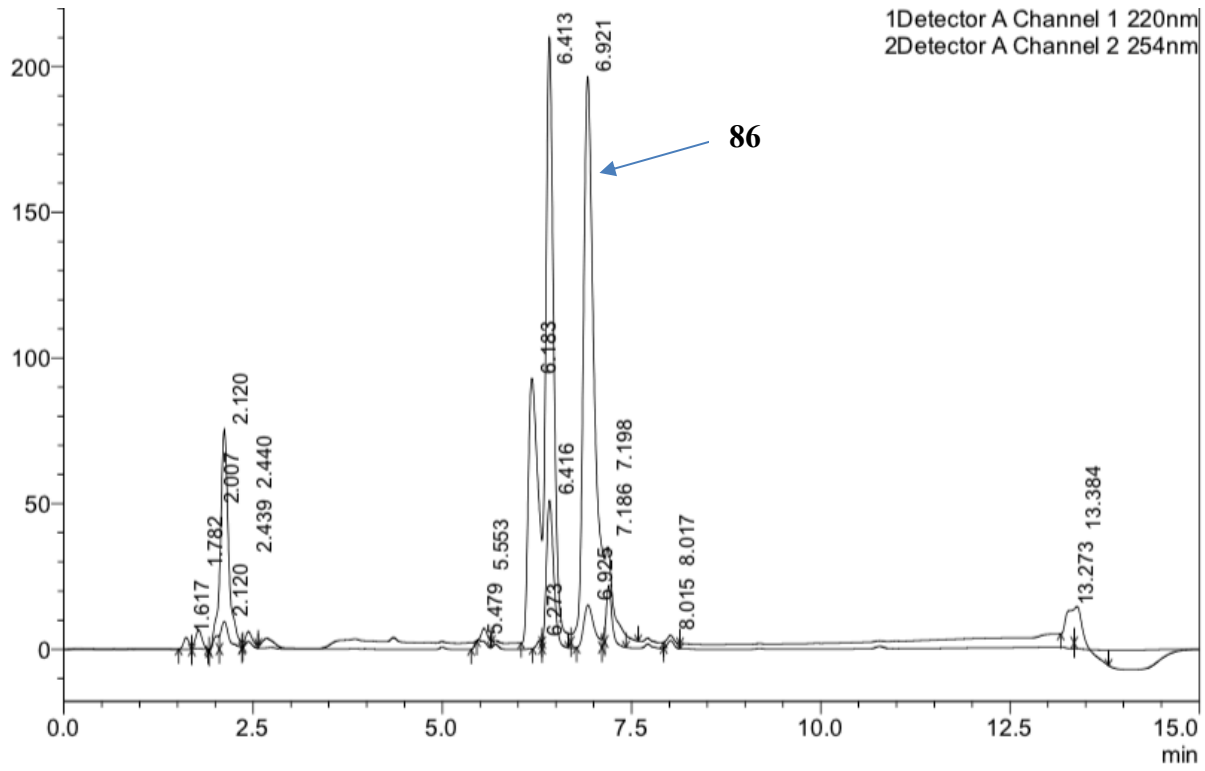


Figure 43 HPLC trace of the crude reaction mixture of exenatide•OAc and **77**

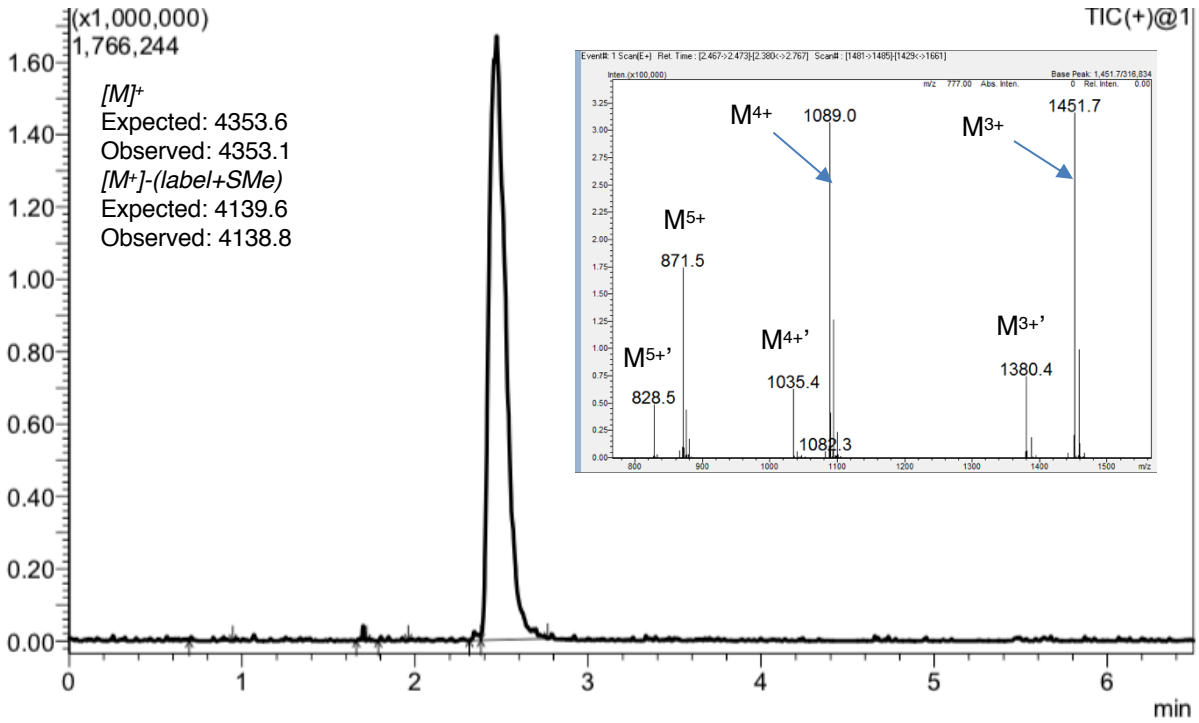
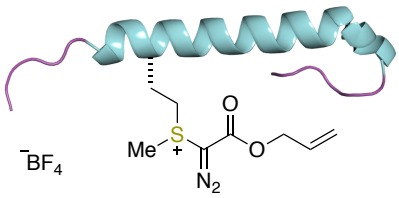


Figure 44 Mass trace of purified exenatide conjugate **86**, with mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with ‘) giving: *M+label-(label+SMe) = M+167-214*.

Exenatide acetate functionalization with **78** *to form* **87**



The labelling of exenatide was performed at room temperature using General Procedure D using **78** (100 mM in CD_3CN) and 1:1 diethyl ether:ethyl acetate for the extraction. The resulting solution was then analysed directly via LC/MS (C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 mL/min) and was observed to have proceeded in >95% conversion.

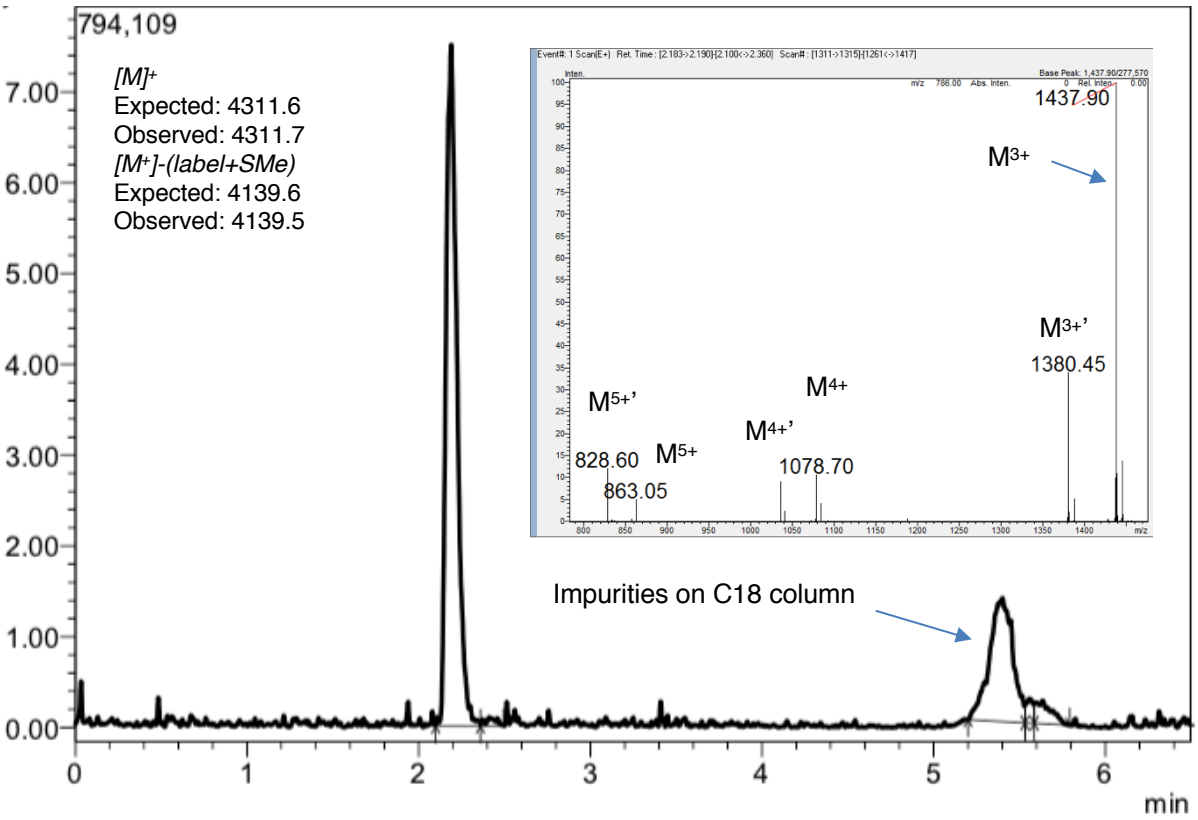
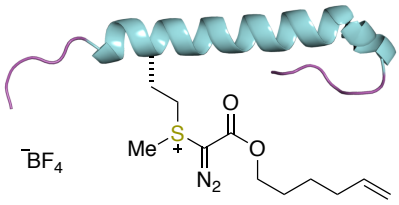


Figure 45 Mass trace of crude reaction mixture of exenatide with **78**, mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with ') giving: *M+label-(label+SMe) = M+125-172*.

Exenatide acetate functionalization with **79** *to form* **88**



The labelling of exenatide was performed at room temperature using General Procedure D using **79** (100 mM in CH_3CN) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) for the extraction. The resulting solution was then analysed directly via LC/MS (C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 mL/min) and was observed to have proceeded in ~90% conversion.

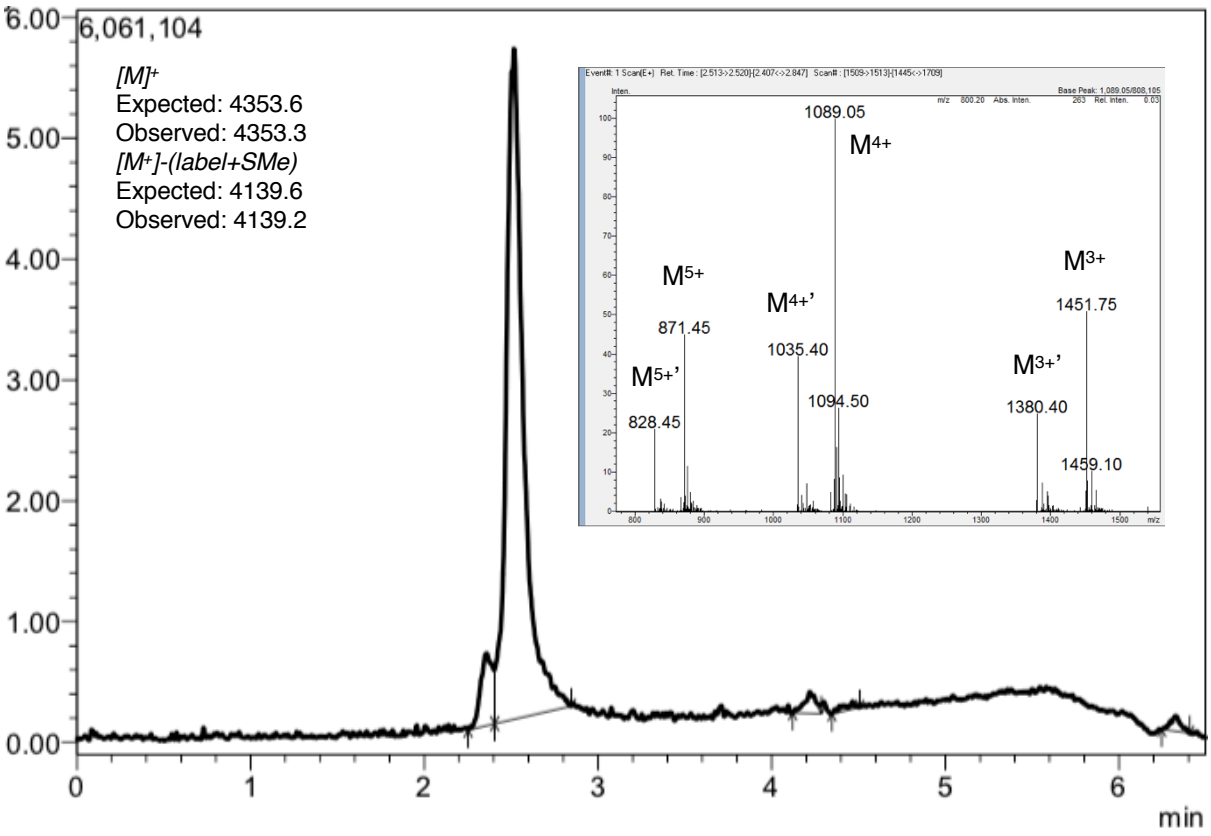
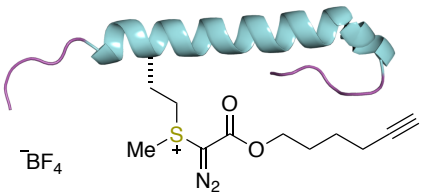


Figure 46 Mass trace of crude reaction mixture of exenatide with **79**, with mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with ') giving: *M+label-(label+SMe) = M+(167)-214*.

Exenatide acetate functionalization with **80** *to form* **89**



The labelling of exenatide was performed at room temperature using General Procedure D using **80** (100 mM in CH_3CN) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) for the extraction. The resulting solution was then analysed directly via LC/MS (C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 mL/min) and was observed to have proceeded in >95% conversion.

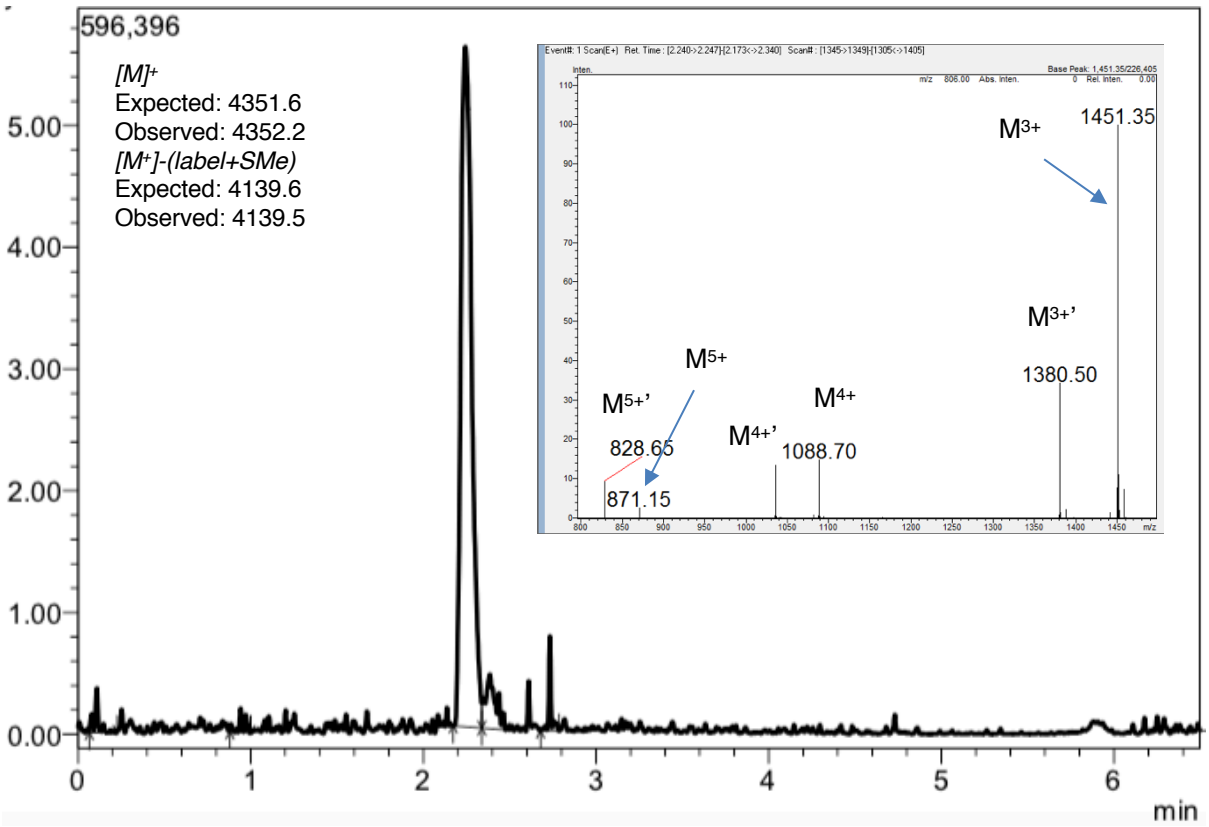
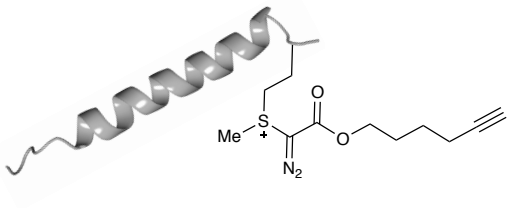


Figure 47 Mass trace of crude reaction mixture of exenatide with **80**, mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with ‘) giving: *M+label-(label+SMe) = M+(165)-212*.

Glucagon functionalization with **80**



The labelling of glucagon was performed at room temperature using General Procedure D using **80** (100 mM in CH_3CN) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) for the extraction. The resulting solution was then analysed directly via LC/MS (C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 mL/min) and was observed to have proceeded in >95% conversion.

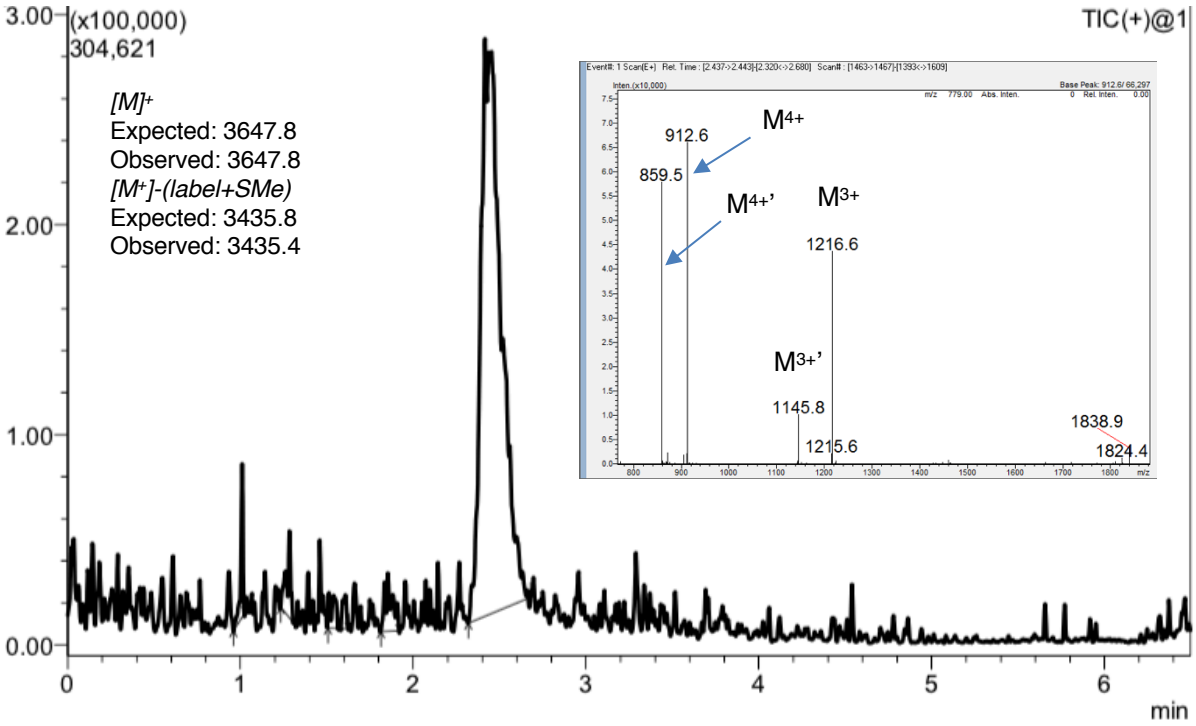
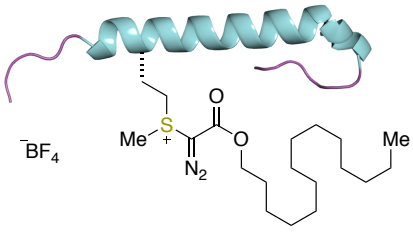


Figure 48 Mass trace of crude reaction mixture of glucagon with **80**, mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with ‘) giving: *M+label-(label+SMe) = M+165-212*.

Exenatide acetate functionalization with **81** *to form* **90**



The labelling of exenatide was performed at room temperature using General Procedure D using **81** (100 mM in CH_3CN) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) for the extraction. The resulting solution was then analysed directly via LC/MS (C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 mL/min) and was observed to have proceeded in ~90% conversion.

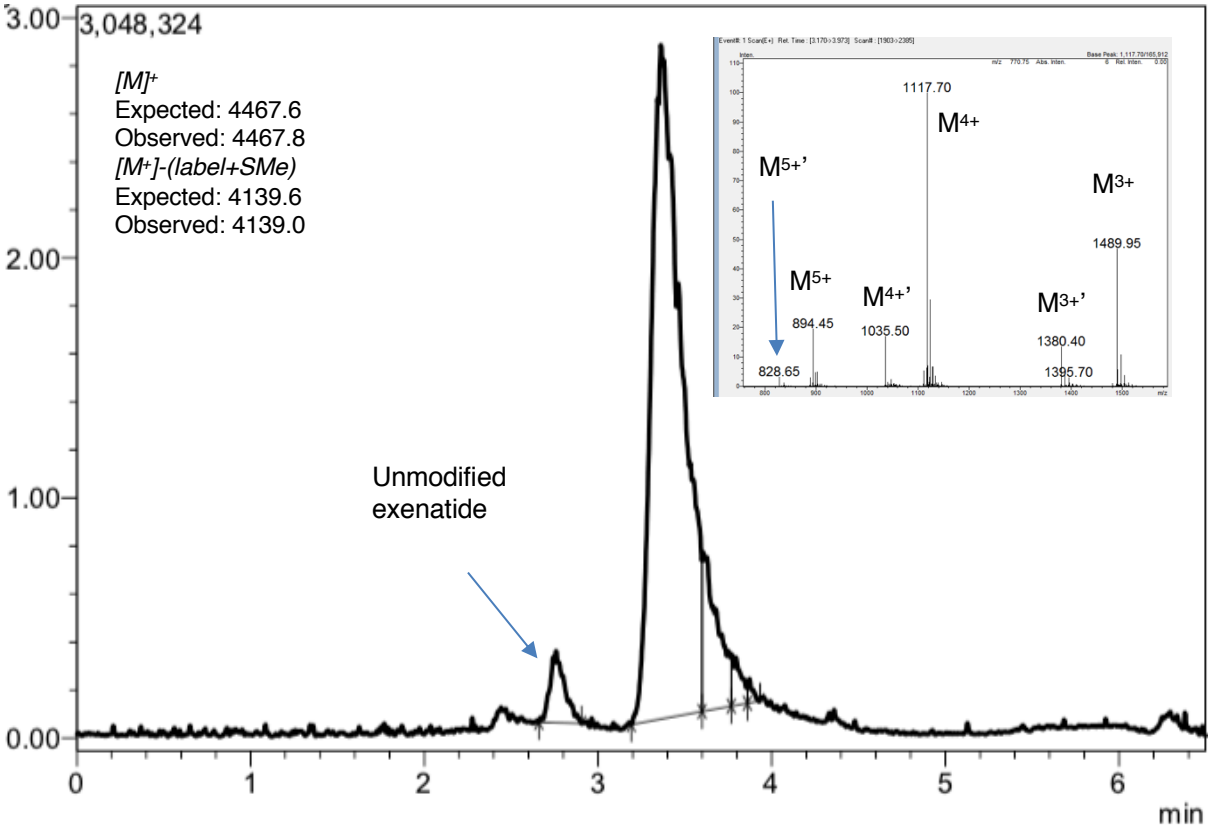
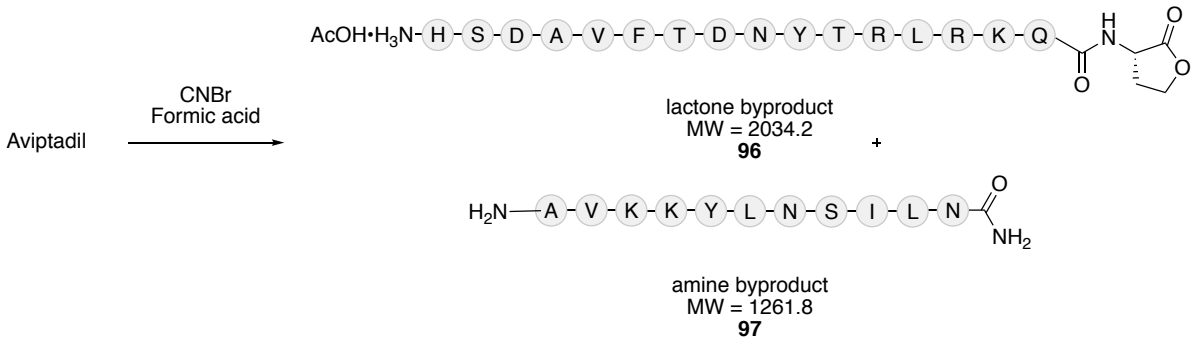


Figure 49 Mass trace of crude reaction mixture of exenatide with **81**, mass chromatogram shown as an insert. Diagnostic fragmentation of C–S bond (indicated with ') giving: *M+label-(label+SMe) = M+(281)-328*.

5.4 Cyanogen bromide mediated backbone cleavage of Aviptadil

Aviptadil (1.1 mg, 0.33 µmol) was added to a 2 mL vial equipped with a magnetic stirrer. To the vial was added a solution of cyanogen bromide (8.5 mg, 83 µmol) in formic acid (200 µL). The resulting solution was stirred at room temperature for 2 hours. The solution was diluted with water then analysed directly via LC/MS (C18 column (50 × 2.1 mm, 2.6µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min), and was observed to have proceeded in >95% conversion to cleavage products **96** and **97**.



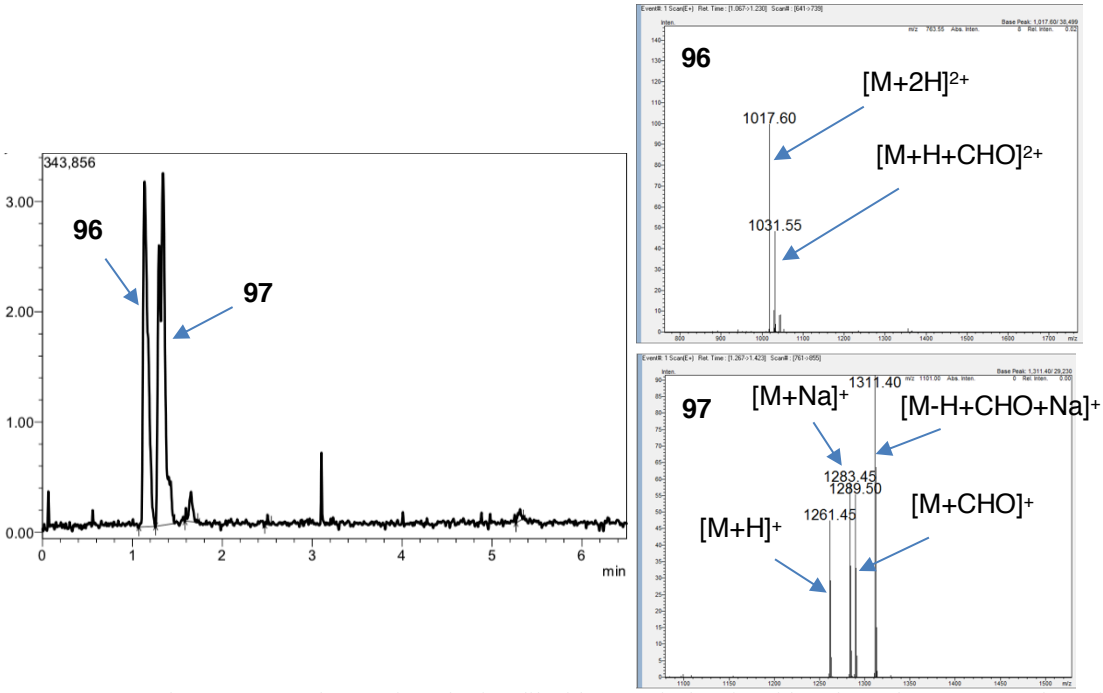
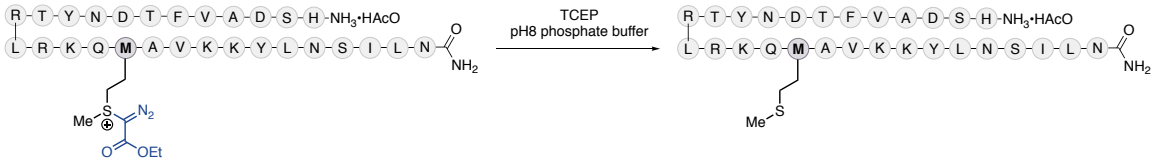


Figure 50 Crude LCMS trace for reaction of aviptadil with CNBr in formic acid. Left; crude mass trace of reaction, right; mass spectra of both peaks corresponding to cleaved polypeptide fragments **96** and **97**.

5.5 Experimental procedures for exploitation of diazo motif in β-sulfonium α-diazoester conjugates

5.5.1 Experimental procedures for phosphine promoted cleavage of methionine label

Phosphine promoted "deligation" of **58**



A vial was charged with a solution of sulfonium conjugate **58** (750 µM, 25 µL). To the vial was added tris-(2-carboxyethyl) phosphine (1.7 mM in pH 8 phosphate buffer, 50 µL) and H_2O (25 µL). The vial was stirred vigorously for 30 minutes. The resulting solution was then analysed directly via LC/MS (C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 mL/min). The deligation was observed to have proceeded in >95% conversion.

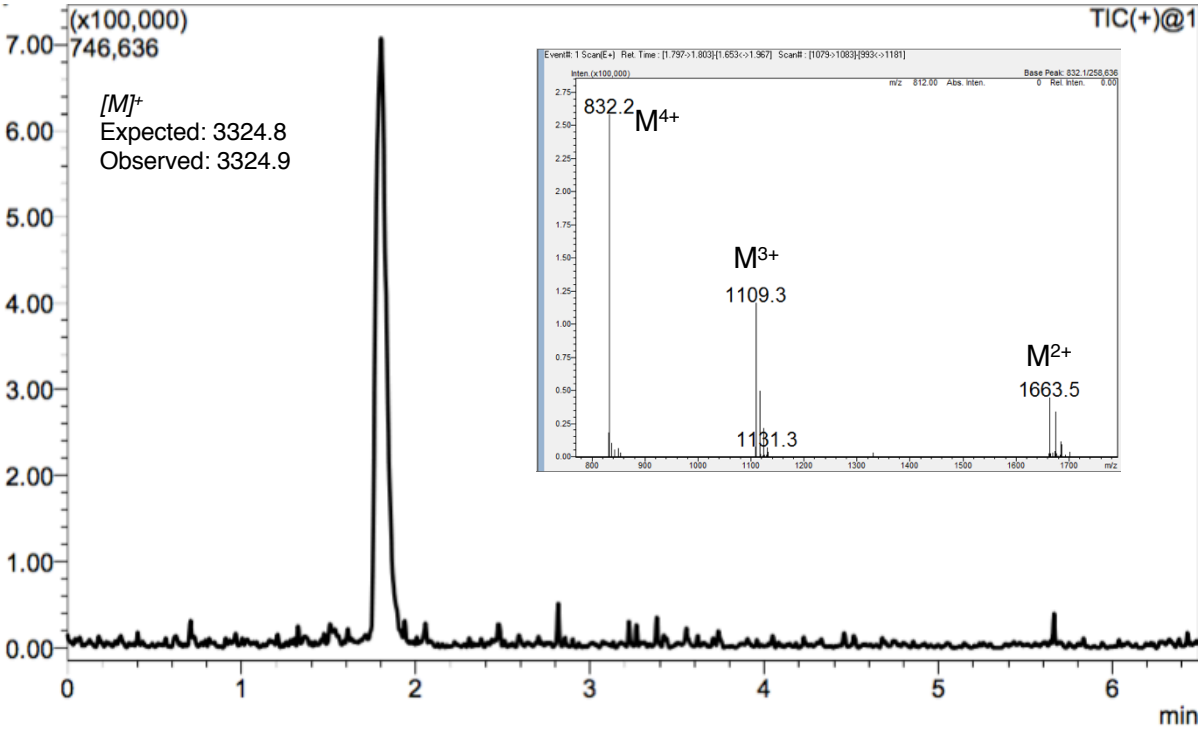
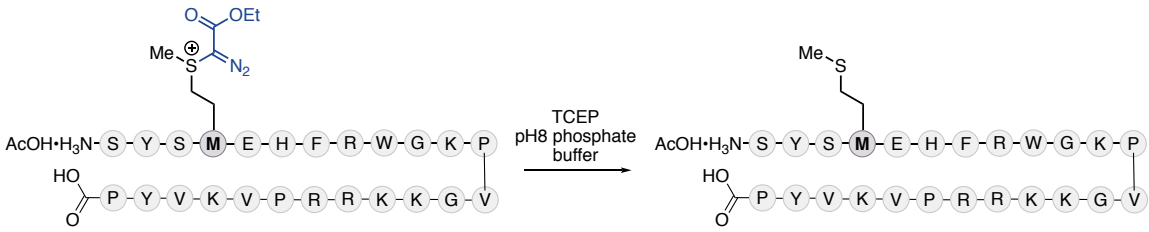


Figure 51 Mass trace of crude reaction mixture of aviptadil conjugate **58** with TCEP at pH 8, mass chromatogram shown as an insert.

Phosphine promoted "deligation" of **57**



A vial was charged with a solution of sulfonium conjugate **57** (650 µM, 15 µL). To this was added tris-(2-carboxyethyl) phosphine (1.7 mM in pH 8 phosphate buffer, 50 µL) and H_2O (35 µL). The vial was stirred vigorously for 30 minutes. The resulting solution was then analysed directly via LC/MS (C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 mL/min), and the ligation reversal was observed to have proceeded in >95% conversion.

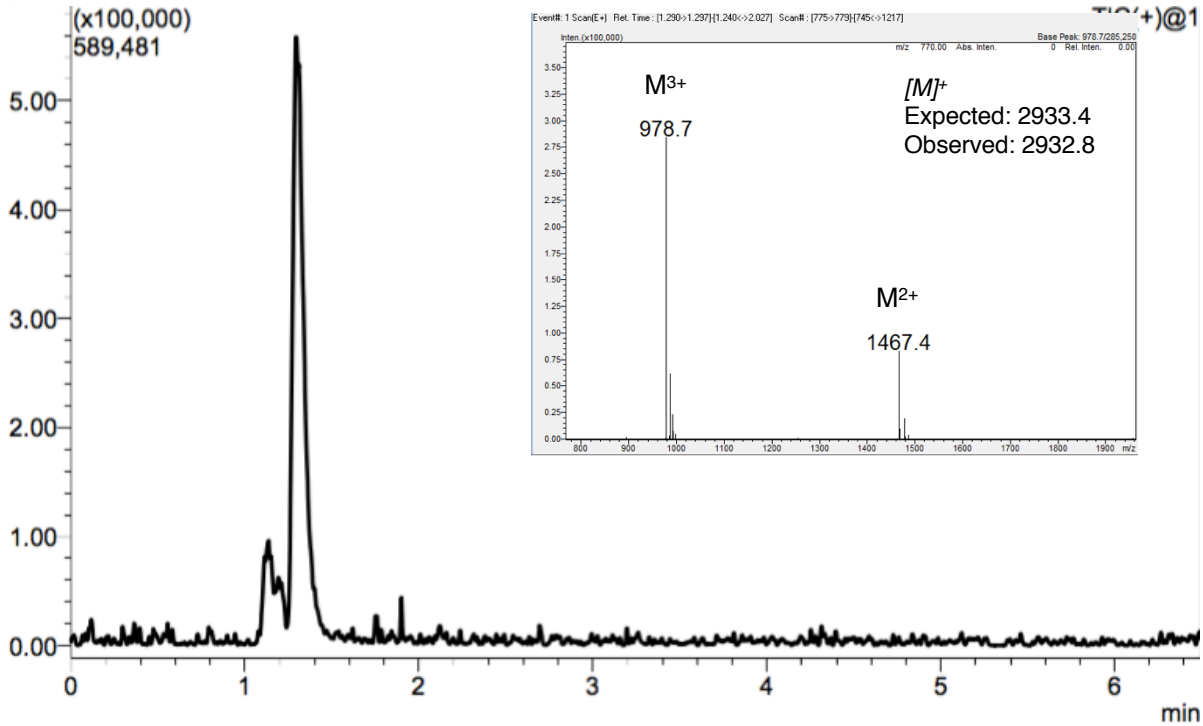
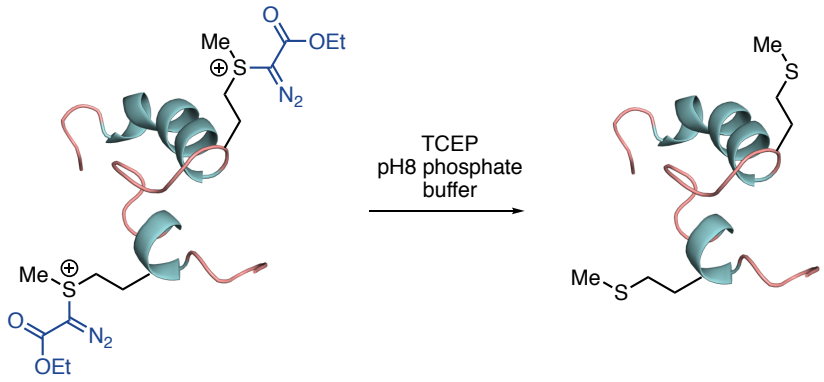


Figure 52 Mass trace of crude reaction mixture of tetracosactide conjugate **57** with TCEP at pH 8, mass chromatogram shown as an insert.

Phosphine promoted "deligation" of **59**



A vial was charged with a solution of sulfonium conjugate **59** (500 µM, 20 µL). To this was added a solution of tris-(2-carboxyethyl) phosphine (1.7 mM in pH 8 phosphate buffer, 50 µL) and H_2O (30 µL). The vial was stirred vigorously for 30 minutes. The resulting solution was then analysed directly via LC/MS (C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 mL/min), and the ligation reversal was observed to have proceeded in >95% conversion.

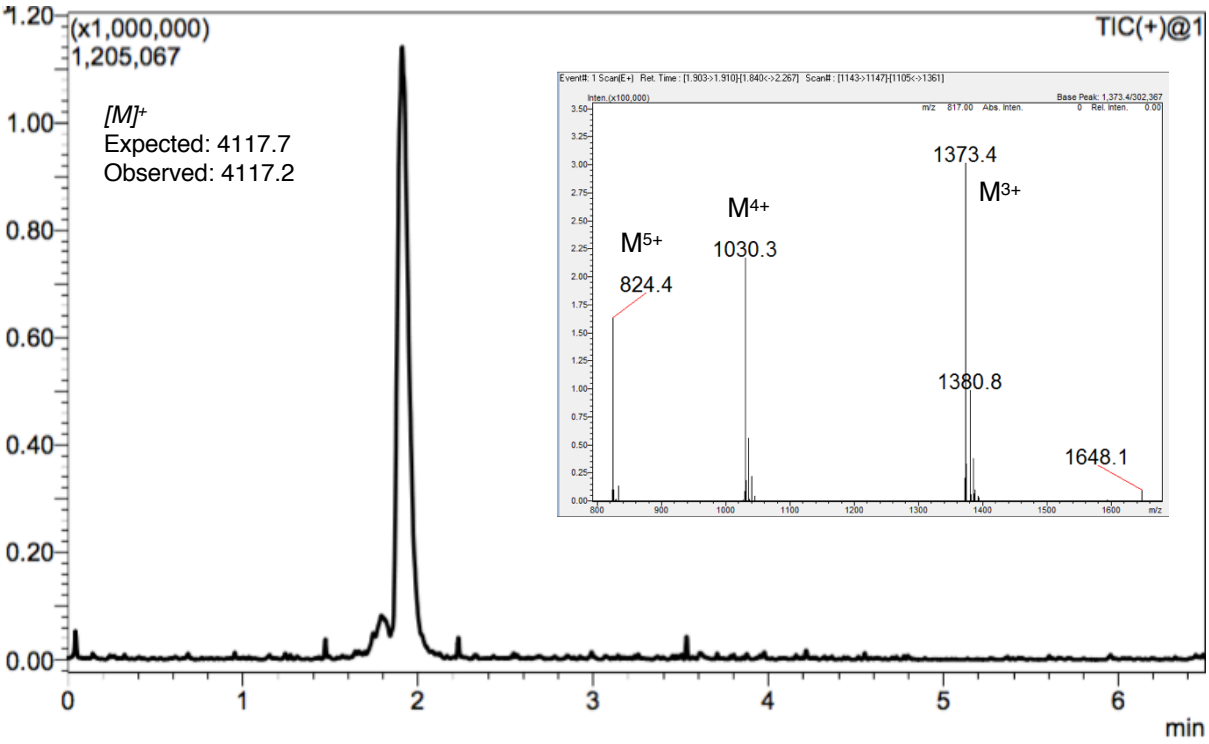
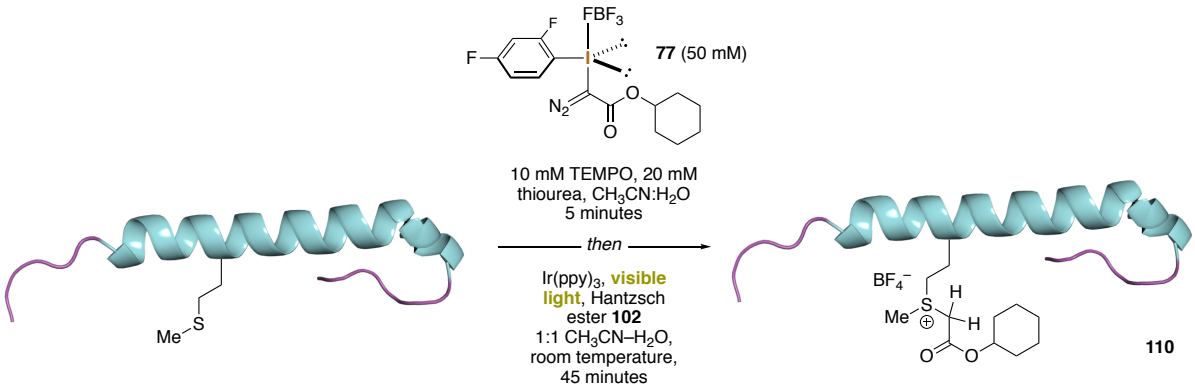


Figure 53 Mass trace of crude reaction mixture of teripartide conjugate **59** with TCEP at pH 8, mass chromatogram shown as an insert.

5.5.2 *Experimental procedures for photochemical reduction of α-sulfonium diazoesters*

Two-step sequential Methionine-Labelling/photochemical Reduction of Exenatide



A 5 mL microwave tube equipped with a magnetic stirrer was charged with a solution of exenatide•OAc (3.7 mM in 0.1 M thiourea (aq), 70 µL, 1 mg). To the vial was added formic acid (100 mM in H_2O, 20 µL), and the resulting solution chilled in an ice bath before TEMPO (33 mM in H_2O, 50 µL), H_2O (30 µL) and **77** (100 mM in CH_3CN, 170 µL) were added. The reaction mixture was stirred for 5 minutes then extracted with 1:1 ethyl acetate:diethyl ether (3 × 0.5 mL), and the excess volatiles were removed in vacuo. To the resulting solution was added a magnetic stirrer and the Hantzsch ester **102** (1.8 mg, 1.0×10^{-5} mmol). The tube was then capped with a septum and subsequently evacuated/backfilled with N_2 (× 5). To the tube was added added fac-Ir(ppy)$_3$ (560 µM in freshly degassed CH_3CN, 180 µL) and freshly degassed H_2O (180 µL). The mixture was irradiated with a Prolite 30 W fluorescent light bulb that was placed at a 5 cm distance from the microwave tube for 45 minutes. A constant temperature was maintained by placing the microwave tube under a stream of air over the course of the irradiation. The reaction mixture was then extracted with 1:1 ethyl acetate:diethyl ether (3 × 0.5 mL) and the excess volatiles were removed in vacuo. The resulting solution was then analysed directly via LC/MS (C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min), and the ligation was observed to have proceeded in >95% conversion. Purification via reverse-phase HPLC (C18 column (150 × 10 mm, 10µm), Gradient: 5-55% B over 9 min then 55-95%B over 1 min, hold for 1.5 min, then 95-5%B over 0.5 min, hold for 3 min, 5 ml/min), followed by lyophilisation yielded **110** as a free-flowing, white solid that was immediately re-dissolved in H_2O (0.5 mL) and stored at –20°C. Concentration analysis by A_{280} indicated a 1.0 mg/mL solution (50% isolated yield).

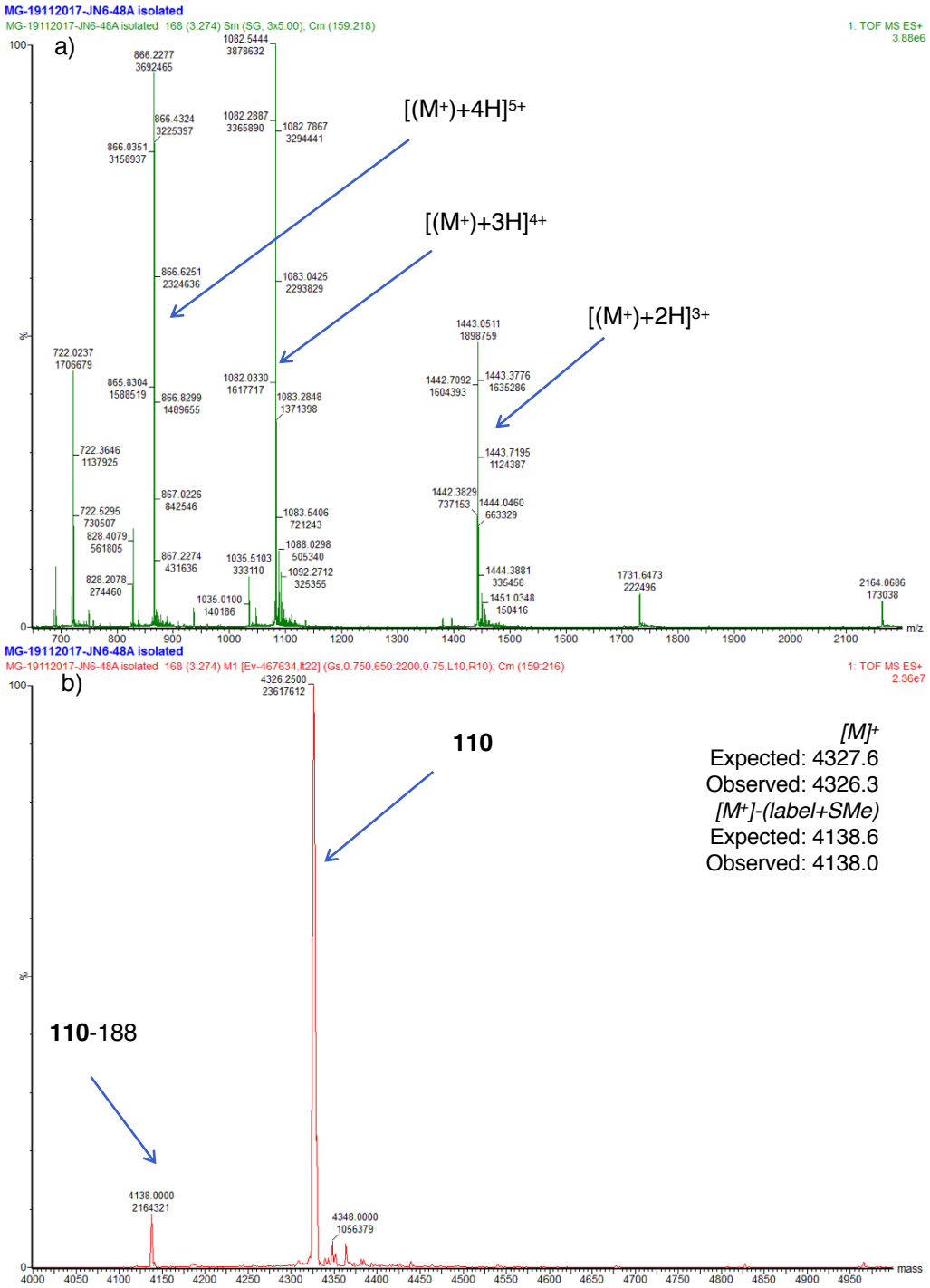
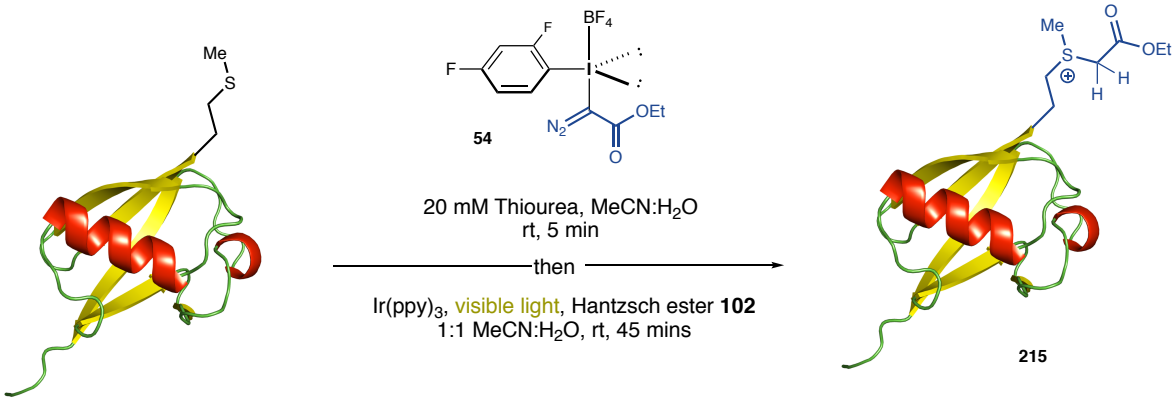


Figure 54 a) Mass spectrum of the crude reaction mixture of the labeling/photoreduction of exenatide with **77** to form **110** b) Deconvoluted mass spectrum of the crude reaction to form **110**

Two-step sequential Methionine-Labelling/photochemical Reduction of Ubiquitin



A 2 mL vial equipped with a magnetic stirrer was charged with Ubiquitin (~1.0 mg, ~0.12 µmol) and evacuated/backfilled with N_2 (× 5). To the vial was added formic acid (100 mM in H_2O, 25 µL) and thiourea (0.1 M in H_2O, 100 µL) and H_2O (125 µL) and the solution stirred vigorously before **54** (100 mM in CH_3CN, 250 µL) was added. The reaction mixture was stirred for 5 minutes then extracted with 1:1 ethyl acetate: diethyl ether (3 × 0.5 mL), and the excess volatiles were removed in vacuo. To the resulting solution was added a magnetic stirrer and the Hantzsch ester (2.0 mg, 1.0×10^{-5} mmol). The tube was then capped with a septum and subsequently evacuated/backfilled with N_2 (× 5). To the tube was added fac-Ir(ppy)$_3$ (560 µM in freshly degassed CH_3CN, 200 µL). The mixture was irradiated with a Prolite 30 W fluorescent light bulb that was placed at a 5 cm distance from the microwave tube for 30 minutes. A constant temperature was maintained by placing the microwave tube under a stream of air over the course of the irradiation. The reaction mixture was then extracted with 1:1 ethyl acetate:diethyl ether (3 × 0.5 mL) and the excess volatiles were removed in vacuo. The resulting solution was then purified directly using MonoS 4.6/100 PE column (GE Healthcare) using a NaCl gradient (0-1 M) in 50 mM ammonium acetate pH 4.5.

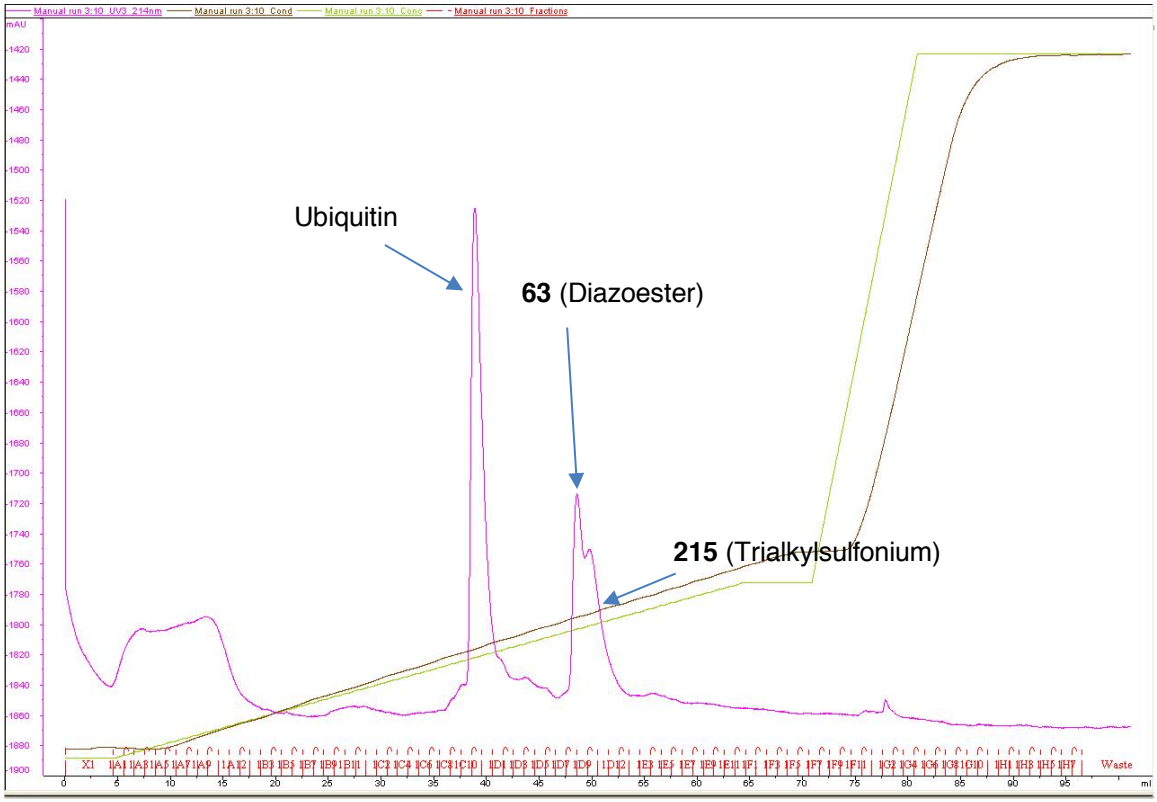


Figure 55 FPLC trace of Ubiquitin conjugate **215**, purified using MonoS 4.6/100 PE column (GE Healthcare) using a NaCl gradient (0-1 M) in 50 mM ammonium acetate pH 4.5. Clear separation of unmodified Ubiquitin and modified conjugates **63** and **215** was achieved due to the differing elution times of the charged species and uncharged species, unmodified ubiquitin using an ion exchange column.

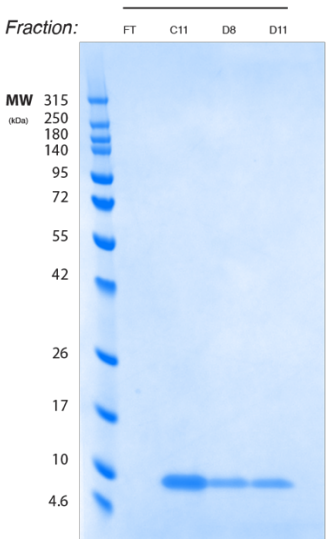
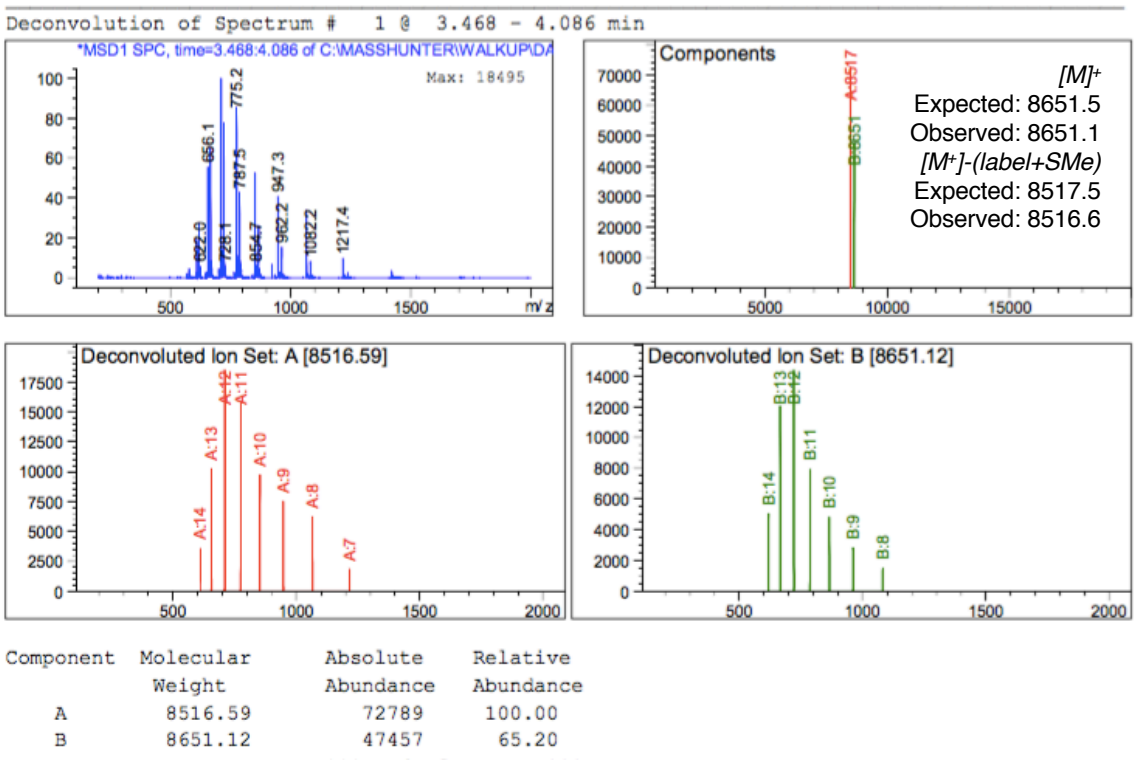
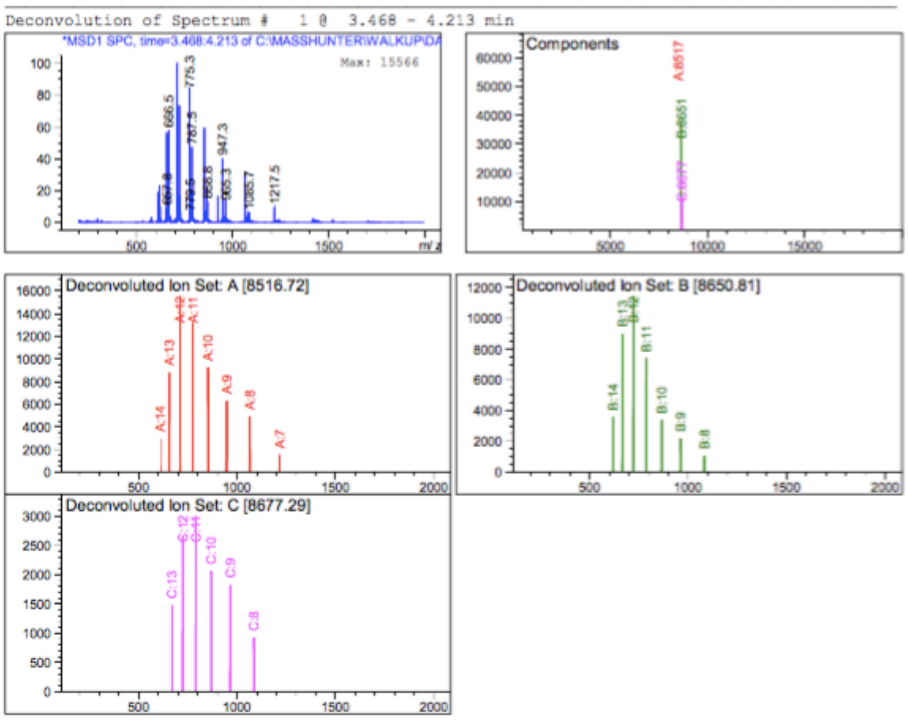


Figure 56 Gel of purified Ubiquitin fractions from FPLC purification run with Bolt 4-12% Bis-Tris gels (Thermo Scientific) and were run for 30 min at 165V (constant) using MES buffer (Thermo Scientific). The ladder on the gel is ProSieve QuadColor protein marker (4.6 kDa – 300 kDa; Lonza) and the gel was stained using InstantBlue (Generon)



Component	Molecular Weight	Absolute Abundance	Relative Abundance
A	8516.59	72789	100.00
B	8651.12	47457	65.20

Figure 57 Protein LCMS trace and deconvolution of purified **215** showing only two species present, the total ion of product **215** (8651 Da) and the fragment ion for **215**-134 (8517 Da).



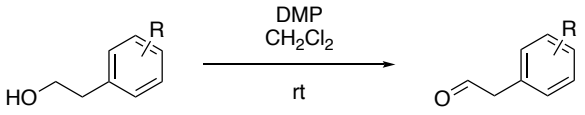
Component	Molecular Weight	Absolute Abundance	Relative Abundance
A	8516.72	61187	100.00
B	8650.81	37658	61.55
C	8677.29	11755	19.21

Figure 58 Protein LCMS trace of mixed fraction of **215** and **63** showing the total ion of product **215** (8651 Da) and the total ion for **63** (8677). Additionally, the same fragment ion for **215**-134 and **63**-160 (8517 Da) is observed

5.5.3 Experimental procedures for a photoredox-mediated bioorthogonal transformation at β-sulfonium α-diazoesters

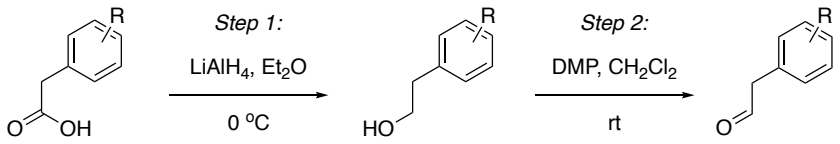
5.5.3.1 Experimental procedures for the synthesis of Hantzsch ester reagents

***General Procedure F**: Synthesis of aldehyde from alcohol*



A solution of alcohol (1.0 eq.) in CH_2Cl_2 (0.1 M) was added to a stirring solution of Dess-Martin Periodinane (1.1 eq.) in CH_2Cl_2 (0.1 M) and the resulting solution stirred under air for 3 hours at room temperature. The reaction was quenched with $Na_2S_2O_3$ (sat. aq.) and the organic layer washed with $NaHCO_3$ (sat. aq.), $Na_2S_2O_3$ (sat. aq.) (x 2) and H_2O. The crude solid was flushed through a plug of silica (50% EtOAc/petroleum ether) and used directly.

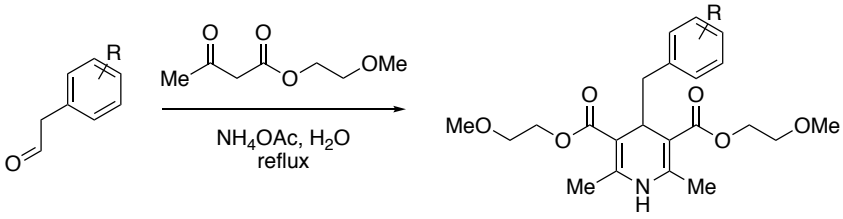
***General Procedure G**: Synthesis of aldehyde from carboxylic acid*



Step 1: A solution of carboxylic acid (1.0 eq.) in Et_2O (0.15 M) was added dropwise to a stirring solution of $LiAlH_4$ (1.5 eq.) in Et_2O (0.15 M) at 0 °C, under N_2. The resulting suspension was allowed to warm to room temperature and stirred for 4 hours. The reaction was quenched with $Na_2SO_4 \cdot 10H_2O$ and stirred for a further 10 minutes. The residue was filtered, and the filtrate evaporated and used directly without further purification.

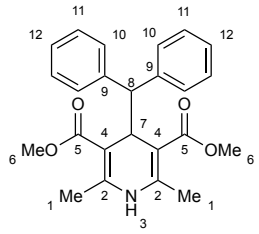
Step 2: A solution of alcohol from step 1 (1.0 eq.) in CH_2Cl_2 (0.1 M) was added to a stirring solution of Dess-Martin Periodinane (1.1 eq.) in CH_2Cl_2 (0.1 M) and the resulting solution stirred under air for 3 hours at room temperature. The reaction was quenched with $Na_2S_2O_3$ (sat. aq.) and the organic layer washed with $NaHCO_3$ (sat. aq.), $Na_2S_2O_3$ (sat. aq.) (× 2) and H_2O. The crude solid was flushed through a plug of silica (50% EtOAc/petroleum ether) and used directly without further purification.

***General Procedure H**: Synthesis of Hantzsch ester from aldehyde*



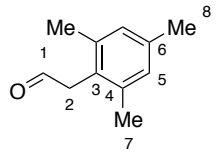
A solution of aldehyde (1 eq.), 2-methoxyethylacetoacetate (4 eq.) and ammonium acetate (2 eq.) in water (0.5 M in aldehyde) was stirred at reflux under air for 4 hours then allowed to cool to room temperature. The reaction mixture was diluted with EtOAc and the aqueous layer extracted with EtOAc (× 3). The combined organic layers were washed with brine, dried over $MgSO_4$, filtered and evaporated to yield the crude product. The crude product was purified by flash column chromatography on silica gel to yield the desired Hantzsch ester.

Dimethyl 4-benzhydryl-2,6-dimethyl-1,4-dihydropyridine-3,5-dicarboxylate **128**



Diphenyl acetaldehyde (0.89 mL, 5.0 mmol) and methyl-3-aminocrotonate (1.15 g, 10.0 mmol) in AcOH (5 mL) were added to a microwave tube and sealed. The solution was degassed with N_2 for 15 minutes then stirred at 110 °C for 5 hours before allowing the solution to cool to room temperature. The reaction mixture was extracted with EtOAc (3 × 5 mL) and the combined organic layers carefully quenched with $NaHCO_3$ (sat. aq.) (15 mL), washed with brine (25 mL), dried over $MgSO_4$, filtered and evaporated to afford the crude product. Purification *via* flash column chromatography on silica gel (50% EtOAc/petroleum ether) and recrystallization from EtOAc/petroleum ether yielded **128** (0.72 g, 37%) as a pale yellow solid. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 7.31 (d, *J* = 7.5 Hz, 4H, H_{10}), 7.21 (t, *J* = 7.7 Hz, 4H, H_{11}), 7.13 (t, *J* = 7.3 Hz, 2H, H_{12}), 5.78 (s, 1H, H_3), 4.86 (d, *J* = 8.9 Hz, 1H, H_7), 3.66 (d, *J* = 8.9 Hz, 1H, H_8), 3.41 (s, 6H, H_6), 2.23 (s, 6H, H_1); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 168.3 (C_5), 144.6 (C_2), 141.5 (C_9), 129.7 (C_{10}), 127.6 (C_{11}), 126.3 (C_{12}), 102.3 (C_4), 58.0 (C_8), 50.7 (C_6), 39.5 (C_7), 19.3 (C_1); **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 392.1859, Calc'd $C_{24}H_{26}O_4N$ requires 392.1856; **IR** ν_{max}/cm^{-1} (film): 3341, 3081, 3023, 3000, 2948, 2888, 1702, 1661, 1623, 1487, 1429, 1213, 1118, 1110, 763, 698; **m.p.**: 172 - 174 °C (decomp)

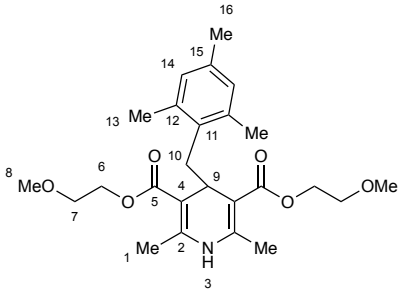
2-Mesitylacetaldehyde **142**



2-mesitylacetaldehyde **142** was synthesised according to General Procedure G, using 2-mesitylacetic acid (0.71 g, 4.0 mmol), LiAlH4 (0.23 g, 6.0 mmol) and Et_2O (13 mL) in Step 1. Dess-Martin periodinane (1.86 g, 4.4 mmol) and CH_2Cl_2 (40 mL) were used in Step 2. **142** was isolated as a colourless oil (0.47 g, 73%). 1H-NMR (400 MHz, CDCl3) δ: 9.66 (t, J = 2.1 Hz,

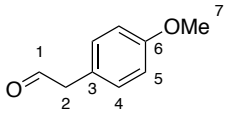
1H, H_1), 6.92 (s, 2H, H_5), 3.73 (d, J = 2.1 Hz, 2H, H_2), 2.29 (s, 3H, H_8), 2.26 (s, 6H, H_7); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 199.2 (C_1), 137.3 (C_4), 137.1 (C_6), 129.3 (C_5), 126.4 (C_3), 44.9 (C_2), 21.0 (C_8), 20.5 (C_7); **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 163.1121, Calc'd $C_{11}H_{15}O^+$ requires 163.1123; **IR** ν_{max}/cm^{-1} (film) 2945, 1720, 1486, 1011

Bis(2-methoxyethyl)2,6-dimethyl-4-(2,4,6-trimethylbenzyl)-1,4-dihydropyridine-3,5-dicarboxylate **148**



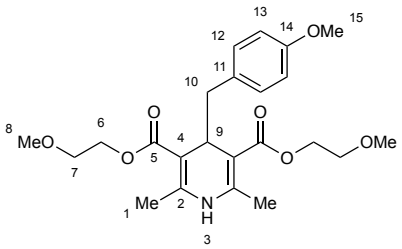
148 was synthesised according to General Procedure H, using **142** (0.32 g, 2 mmol), 2-methoxyethyl acetoacetate (1.17 g, 8.0 mmol) and ammonium acetate (0.31 g, 4.0 mmol) in water (4 mL). Purification by flash column chromatography on silica gel (10-50% EtOAc/40-60 petroleum ether) followed by recrystallization from diethyl ether/EtOAc/hexane afforded **148** (0.16 g, 18%) as an off-white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 6.75 (s, 2H, H_{14}), 5.71 (br. s, 1H, H_3), 4.34 (t, *J* = 7.9 Hz, 1H, H_9), 4.06 (dt, *J* = 4.8, 12.0 Hz, 2H, H_6), 3.82 (dt, *J* = 4.8, 12.0 Hz, 2H, H_6), 3.48 (t, *J* = 4.8 Hz, 4H, H_7), 3.36 (s, 6H, H_8), 2.66 (d, *J* = 7.9 Hz, 2H, H_{10}), 2.31 – 2.28 (m, 12H, H_1 and H_{13}), 2.19 (s, 3H, H_{16}); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 167.8 (C_5), 145.3 (C_2), 137.8 (C_{11}), 135.0 (C_{15}), 132.2 (C_{12}), 128.7 (C_{14}), 103.4 (C_4), 70.6 (C_7), 62.7 (C_6), 59.0 (C_8), 35.2 (C_{10}), 32.7 (C_9), 20.9 (C_{16}), 20.1 (C_{13}), 19.6 (C_1); **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 446.2515, Calc'd $C_{25}H_{36}O_6N$ requires 446.2537; **IR** νmax/cm-1 (film): 3303, 2912 (br), 2806, 1701, 1690, 1619, 1491, 1305, 1266, 1198, 1084, 1058, 1017, 847, 773, 766, 721; **m.p.**: 110 °C (decomp).

2-(4-Methoxyphenyl)acetaldehyde **139**



2-(4-methoxyphenyl)acetaldehyde **139** was synthesised according to General Prodedure F, using 2-(4-methoxyphenyl)ethan-1-ol (2.28 g, 15.0 mmol), Dess-Martin periodinane (7.00 g, 16.5 mmol) and CH_2Cl_2 (150 mL). **139** was isolated as a colourless oil (0.98 g, 43%). **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 9.71 (t, *J* = 2.5 Hz, 1H, H_1), 7.15 – 7.11 (m, 2H, H_4), 6.93 – 6.88 (m, 2H, H_5), 3.80 (s, 3H, H_7), 3.62 (d, *J* = 2.5 Hz, 2H, H_2); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 199.8 (C_1), 159.0 (C_6), 130.8 (C_4), 123.8 (C_3), 114.6 (C_5), 55.4 (C_7), 49.8 (C_2); **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 151.0759, Calc'd $C_9H_{11}O_2^+$ requires 151.0759; **IR** ν_{max}/cm^{-1} (film): 2934, 1721, 1612, 1513, 1247.

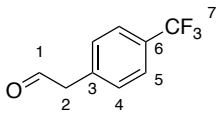
Bis(2-methoxyethyl)4-(4-methoxybenzyl)-2,6-dimethyl-1,4-dihydropyridine-3,5-dicarboxylate **145**



145 was prepared according to General Procedure H using **139** (0.98 g, 6.5 mmol), 2-methoxyethyl acetoacetate (3.8 g, 26.0 mmol) and ammonium acetate (1.0 g, 13.0 mmol) in water (13 mL). Purification by flash column chromatography on silca gel (20-50% EtOAc/40-60 petroleum ether) followed by recrystallization from diethyl ether/hexane afforded **145** (0.97 g, 34%) as an off-white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 6.94 (d, *J* = 8.3 Hz, 2H, H_{12}), 6.72 (d, *J* = 8.3 Hz, 2H, H_{13}), 5.18 (br. s, 1H, H_3), 4.23 – 4.12 (m, 5H, H_9 and H_6), 3.76 (s, 3H, H_{15}), 3.61 – 3.54 (m, 4H, H_7), 3.40 (s, 6H, H_8), 2.55 (d, *J* = 5.4 Hz, 2H, H_{10}), 2.16 (s, 6H, H_1); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 167.8 (C_5), 158.0 (C_{14}), 145.8 (C_2), 131.56 (C_{11}), 131.2 (C_{12}), 112.9 (C_{13}), 101.8 (C_4), 70.8 (C_7), 62.8 (C_6), 59.1 (C_8), 55.4 (C_{15}), 41.4 (C_{10}), 35.7 (C_9), 19.4 (C_1); **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 434.2176, Calc'd $C_{23}H_{32}O_7N$ requires 434.2173; **IR**

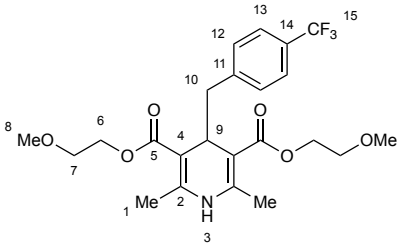
ν_{max}/cm^{-1} (film): 3337, 2985, 2941, 2862, 2885, 2838, 1684, 1653, 1625, 1609, 1497, 1470, 1445, 1224, 1212, 1089, 1029, 856, 796, 697; **m.p.**: 62 - 64 °C (decomp).

2-(4-(Trifluoromethyl)phenyl)acetaldehyde **140**



2-(4-(trifluoromethyl)phenyl)acetaldehyde **140** was synthesised according to General Procedure G, using 2-(4-(trifluoromethyl)phenyl)acetic acid (3.06 g, 15.0 mmol), $LiAlH_4$ (0.85 g, 22.5 mmol) and Et_2O (50 mL) in Step 1 and Dess-Martin periodinane (7.00 g, 16.5 mmol) and CH_2Cl_2 (150 mL) in Step 2. **140** was isolated as a colourless oil (2.03 g) which was used directly without further purification.

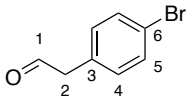
Bis(2-methoxyethyl)2,6-dimethyl-4-(4-(trifluoromethyl)benzyl)-1,4-dihydropyridine-3,5-dicarboxylate **146**



146 was prepared according to General Procedure H using crude aldehyde **140** (0.75 g, 4.0 mmol), 2-methoxyethyl acetoacetate (2.34 g, 16.0 mmol) and ammonium acetate (0.62 g, 8.0 mmol) in water (13 mL). Purification by flash column chromatography on silca gel (10-50% EtOAc/40-60 petroleum ether) followed by recrystallization from diethyl ether/hexane afforded **146** (0.60 g, 32%) as an off-white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 7.43 (d, J = 7.9 Hz, 2H, H_{13}), 7.16 (d, J = 7.9 Hz, 2H, H_{12}), 5.33 (br. s, 1H, H_3), 4.25 (t, J = 5.6 Hz, 1H, H_9), 4.21 – 4.07 (m, 4H, H_6), 3.60 – 3.50 (m, 4H, H_7), 3.39 (s, 6H, H_8), 2.67 (d, J = 5.6 Hz, 2H, H_{10}), 2.18 (s, 6H, H_1); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 167.5 (C_5), 146.0 (C_2), 143.9 (C_{11}), 130.5 (C_{12}), 128.1 (q, $^2J_{C–F}$ = 32.6 Hz, C_{14}), 124.6 (q, $^1J_{C–F}$ = 270.2 Hz, C_{15}), 124.2 (C_{13}), 101.6 (C_4), 70.8

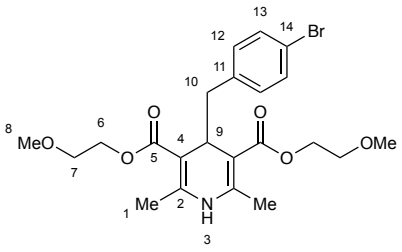
(C_7), 62.9 (C_6), 59.1 (C_8), 42.4 (C_{10}), 35.7 (C_9), 19.4 (C_1); **^{19}F{H}-NMR** (376 MHz, $CDCl_3$) δ: -64.3; **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 472.1949, Calc'd $C_{23}H_{29}F_3O_6N$ requires 472.1941; **IR** ν_{max}/cm^{-1} (film): 3339, 2934 (br), 2885 (br), 1673, 1652, 1615, 1477, 1323, 1224, 1165, 1099, 1089, 1026, 1015, 859, 773, 709; **m.p.**: 102 - 104 °C (decomp).

2-(4-Bromophenyl)acetaldehyde **141**



2-(4-bromophenyl)acetaldehyde **141** was synthesised according to General Procedure G, using 2-(4-bromophenyl)acetic acid (3.23 g, 15.0 mmol), $LiAlH_4$ (0.85 g, 22.5 mmol) and Et_2O (50 mL) in Step 1 and Dess-Martin periodinane (7.00 g, 16.5 mmol) and CH_2Cl_2 (150 mL) in Step 2. **141** was isolated as a colourless oil (1.91 g, 64%). **^{1}H-NMR** (400 MHz, CDCl3) δ: 9.74 (t, *J* = 2.1 Hz, 1H, H_1), 7.51 – 7.47 (m, 2H, H_5), 7.11 – 7.07 (m, 2H, H_4), 3.66 (d, *J* = 2.1 Hz, 2H, H_2); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 198.7 (C_1), 132.2 (C_4), 131.4 (C_5), 130.9 (C_3), 121.7 (C_6), 50.0 (C_2); **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 198.9758, Calc'd $C_8H_8{}^{79}BrO^+$ requires 198.9759; **IR** νmax/cm^{-1} (film): 2926, 1720, 1488, 1071.

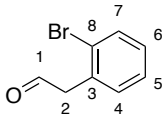
Bis(2-methoxyethyl)4-(4-bromobenzyl)-2,6-dimethyl-1,4-dihydropyridine-3,5-dicarboxylate **147**



147 was prepared according to General Procedure H using **141** (0.80 g, 4.0 mmol), 2-methoxyethyl acetoacetate (2.3 g, 16.0 mmol) and ammonium acetate (0.62 g, 8.0 mmol) in water (8 mL). Purification by flash column chromatography on silica gel (10-50% EtOAc/40-60 petroleum ether) followed by recrystallization from diethyl ether/hexane afforded **147** (0.47

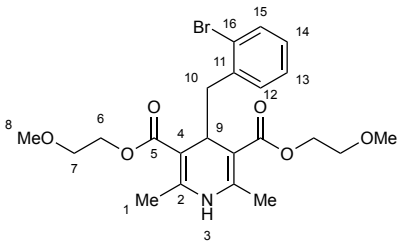
g, 24%) as an off-white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 7.29 (d, *J* = 8.1 Hz, 2H, H_{13}), 6.91 (d, *J* = 8.1 Hz, 2H, H_{12}), 5.23 (br. s, 1H, H_3), 4.23 – 4.15 (m, 5H, H_6 and H_9), 3.62 – 3.51 (m, 4H, H_7), 3.40 (s, 6H, H_8), 2.57 (d, *J* = 5.4 Hz, 2H, H_{10}), 2.17 (s, 6H, H_1); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 167.6 (C_5), 146.0 (C_2), 138.6 (C_{11}), 132.0 (C_{12}), 130.4 (C_{13}), 119.8 (C_{14}), 101.6 (C_4), 70.8 (C_7), 62.8 (C_6), 59.1 (C_8), 41.7 (C_{10}), 35.6 (C_9), 19.5 (C_1); **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 484.1162, Calc'd $C_{22}H_{29}{}^{81}BrO_6N$ requires 484.1152; **IR** ν_{max}/cm^{-1} (film): 3338, 3043, 2981, 2879, 2811, 1671, 1654, 1481, 1373, 1224, 1123, 1024, 1009, 770; **m.p.**: 112 - 114 °C (decomp).

2-(2-Bromophenyl)acetaldehyde **143**



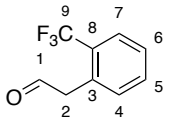
2-(2-bromophenyl)acetaldehyde **143** was synthesised according to General Prodedure F, using 2-(2-bromophenyl)ethan-1-ol (1.36 mL, 10.0 mmol), Dess-Martin periodinane (4.67 g, 11.0 mmol) and CH_2Cl_2 (100 mL). **143** was isolated as a colourless oil (1.21 g, 61%). **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 9.77 – 9.75 (m, 1H, H_1), 7.65 – 7.60 (m, 1H, H_7), 7.35 – 7.29 (m, 1H, H_5), 7.27 – 7.22 (m, 1H, H_4), 7.22 – 7.16 (m, 1H, H_6), 3.88 – 3.86 (m, 2H, H_2); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 198.4 (C_1), 133.2 (C_7), 132.8 (C_3), 131.9 (C_4), 129.4 (C_6), 128.0 (C_5), 125.1 (C_8), 50.6 (C_2); **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 198.9754, Calc'd $C_8H_8{}^{79}BrO^+$ requires 198.9759; **IR** ν_{max}/cm^{-1} (film): 2937, 1723, 1472, 1027.

Bis(2-methoxyethyl)-4-(2-bromobenzyl)-2,6-dimethyl-1,4-dihydropyridine-3,5-dicarboxylate **149**



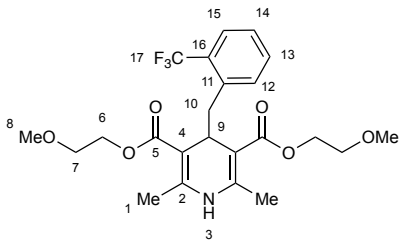
149 was prepared according to General Procedure H using **143** (0.40 g, 2.0 mmol), 2-methoxyethyl acetoacetate (1.17 g, 8.0 mmol) and ammonium acetate (0.31 g, 4.0 mmol) in water (4 mL). Purification by flash column chromatography on silica gel (20-50% EtOAc/40-60 petroleum ether) followed by recrystallization from diethyl ether/hexane afforded **149** (0.36 g, 38%) as an off-white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 7.45 (m, 1H, H_{15}), 7.12 (m, 1H, H_{13}), 7.04 (dd, *J* = 1.6, 7.7 Hz, 1H, H_{12}), 6.99 (td, *J* = 1.6, 7.7 Hz, 1H, H_{14}), 5.56 (br. s, 1H, H_3), 4.38 (t, *J* = 6.1 Hz, 1H, H_9), 4.20 – 4.13 (m, 2H, H_6), 4.05 – 3.99 (m, 2H, H_6), 3.57 (t, *J* = 5.0 Hz, 4H, H_7), 3.39 (s, 6H, H_8), 2.80 (d, *J* = 6.1 Hz, 2H, H_{10}), 2.24 (s, 6H, H_1); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 167.6 (C_5), 146.2 (C_2), 138.8 (C_{11}), 132.7 (C_{12}), 132.3 (C_{15}), 127.5 (C_{14}), 126.4 (C_{13}), 125.8 (C_{16}), 101.8 (C_4), 70.8 (C_7), 62.9 (C_6), 59.1 (C_8), 41.2 (C_{10}), 34.3 (C_9), 19.6 (C_1); **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 484.1140, Calc'd $C_{22}H_{29}{}^{81}BrO_6N$ requires 484.1152; **IR** ν_{max}/cm^{-1} (film): 3337, 2923 (br), 2890, 2812, 1687, 1635, 1616, 1486, 1471, 1263, 1200, 1089, 1024, 760, 751; **m.p.**: 68 - 70 °C (decomp).

2-(2-(Trifluoromethyl)phenyl)acetaldehyde **144**



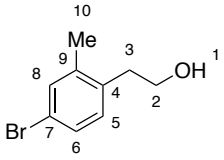
2-(2-(trifluoromethyl)phenyl)acetaldehyde **144** was synthesised according to General Procedure F, using 2-(2-(trifluoromethyl)phenyl)ethan-1-ol (1.59 mL, 10.0 mmol), Dess-Martin periodinane (4.67 g, 11.0 mmol) and CH_2Cl_2 (100 mL). **144** was isolated as a colourless oil (1.17 g), which was used directly without further purification.

Bis(2-methoxyethyl)2,6-dimethyl-4-(2-(trifluoromethyl)benzyl)-1,4-dihydropyridine-3,5-dicarboxylate **150**



150 was prepared according to General Procedure H using **144** (0.38 g, 2.0 mmol), 2-methoxyethyl acetoacetate (1.17 g, 8.0 mmol) and ammonium acetate (0.31 g, 4.0 mmol) in water (4 mL). Purification by flash column chromatography on silica gel (10-50% EtOAc/40-60 petroleum ether) followed by recrystallization from diethyl ether/hexane afforded **150** (0.35 g, 37%) as an off-white solid. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 7.56 (d, J = 7.6 Hz, 1H, H_{15}), 7.38 (t, J = 7.6 Hz, 1H, H_{13}), 7.23 (m, 2H, H_{12} and H_{14}), 5.85 (br. s, 1H, H_3), 4.36 (t, J = 7.3 Hz, 1H, H_9), 4.13 – 4.06 (m, 2H, H_6), 3.91 – 3.84 (m, 2H, H_6), 3.53 – 3.44 (m, 4H, H_7), 3.36 (s, 6H, H_8), 2.76 (d, J = 7.3 Hz, 2H, H_{10}), 2.32 (s, 6H, H_1); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 167.5 (C_5), 146.1 (C_2), 137.9 (C_{11}), 133.5 (C_{12} or C_{14}), 130.8 (C_{13}), 129.14 (q, $^{2}J_{C–F}$ = 30.0 Hz, C_{16}), 125.9 (C_{12} or C_{14}), 125.6 (q, $^{3}J_{C–F}$ = 5.7 Hz, C_{15}), 124.8 (q, $^{1}J_{C–F}$ = 275.1 Hz, C_{17}), 102.5 (C_4), 70.5 (C_7), 62.7 (C_6), 59.0 (C_8), 38.4 (C_{10}), 34.9 (C_9), 19.6 (C_1); **^{19}F{H}-NMR** (376 MHz, $CDCl_3$) δ: -60.9; **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 472.1925, Calc'd $C_{23}H_{29}F_3O_6N$ requires 472.1941; **IR** ν_{max}/cm^{-1} (film): 3327, 2928 (br), 2821, 1691, 1639, 1617, 1489, 1310, 1265, 1206, 1206, 1157, 1112, 1089, 1057, 1032, 771, 744; **m.p.**: 76 °C (decomp)

2-(4-Bromo-2-methylphenyl)ethan-1-ol **161a**

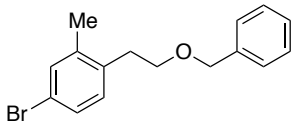


n-Butyl lithium (4.2 mL, 10.5 mmol, 2.5M in hexanes) was added dropwise to a solution of 5-bromo-2-iodotoluene (1.5 mL, 10.5 mmol) in diethyl ether (65 mL) at -78 °C. The resulting mixture was stirred at -78 °C for 10 minutes then ethylene oxide (6.3 mL, 15.8 mmol, 2.5M in THF) was added dropwise. The solution was warmed to room temperature and stirred for a further 2 hours. The reaction was quenched with NH_4Cl (aq.) (30 mL) and extracted with EtOAc (3 × 50 mL). The combined organic layers were washed with water (35 mL) and brine (35 mL), dried over $MgSO_4$, filtered and evaporated to yield the crude product. Purification by flash column chromatography on silica gel (10-30% diethyl ether/40-60 petroleum ether) yielded the title compound (1.43 g, 63%) as a colourless oil. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 7.33 – 7.29 (m, 1H), 7.28 – 7.24 (m, 1H), 7.02 (d, J = 8.1 Hz, 1H), 3.78 (q, J = 6.5 Hz, 2H), 2.82 (t, J = 6.5

Hz, 2H), 2.30 (s, 3H), 1.80 (t, *J* = 5.6 Hz, 1H); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 138.8, 135.7, 133.1, 131.2, 129.0, 120.2, 62.5, 35.8, 19.3.*

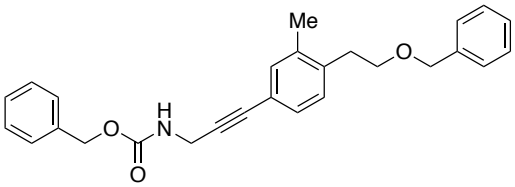
* Data in accordance with the literature[280]

1-(2-(Benzyloxy)ethyl)-4-bromo-2-methylbenzene **161b**



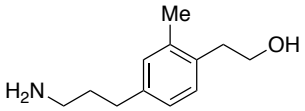
Synthesised by Dr C. He.

Benzyl (3-(4-(2-(benzyloxy)ethyl)-3-methylphenyl)prop-2-yn-1-yl)carbamate **161c**



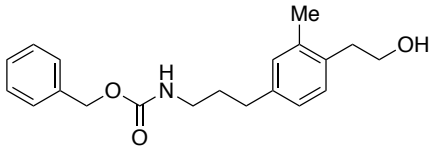
Synthesised by Dr C. He.

2-(4-(3-Aminopropyl)-2-methylphenyl)ethan-1-ol **161d**



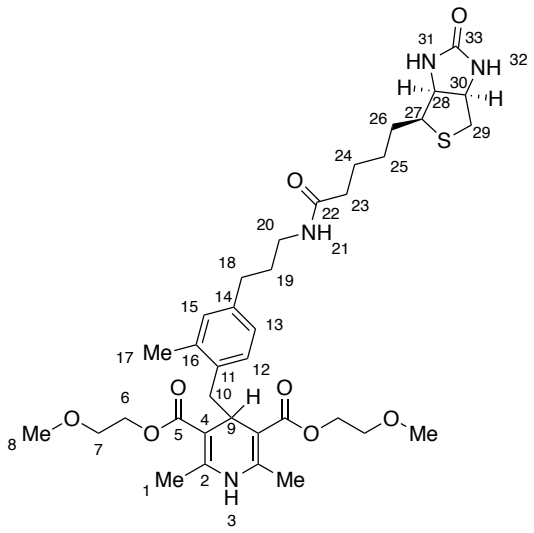
Synthesised by Dr C. He.

Benzyl (3-(4-(2-hydroxyethyl)-3-methylphenyl)propyl)carbamate **161e**



Synthesised by Dr C. He.

Bis(2-methoxyethyl)2,6-dimethyl-4-(2-methyl-4-(3-(5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)propyl)benzyl)-1,4-dihydropyridine-3,5-dicarboxylate **157**



Step 1: A solution of **162** (2.38 g, 7.3 mmol) in CH_2Cl_2 (25 mL) was added dropwise to a solution of Dess-Martin Periodinane (3.39 g, 8.0 mmol) in CH_2Cl_2 (50 mL) and the resulting solution stirred for 4 hours at room temperature. The reaction was quenched with $Na_2S_2O_3$ (sat. aq.) (50 mL) and the organic layer washed with $NaHCO_3$ (sat. aq.) (50 mL), $Na_2S_2O_3$ (sat. aq.) (2 × 50 mL) and H_2O (50 mL). The organic layer was dried over $MgSO_4$, filtered and evaporated to yield the title compound as a crude oil (1.78 g, 76%) which was used directly without further purification.

Step 2: According to General Procedure H, the crude aldehyde **163** (1.71 g, 5.3 mmol), 2-methoxyethyl acetoacetate (3.07 g, 21.0 mmol) and ammonium acetate (0.81 g, 11 mmol) in water (11 mL) were heated to reflux for 7 hours. Purification by flash column chromatography

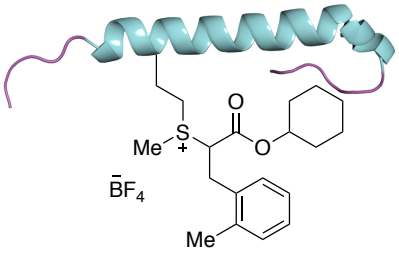
on silica gel (10-50% EtOAc/40-60 petroleum ether) afforded **164** as an impure orange gum which was used directly without further purification.

Step 3: Palladium on carbon (0.36 g, 3.4 mmol, 10 mol% Pd) was added to a solution of **164** (1.86 g, 3.0 mmol) in MeOH (30 mL) under N_2. The flask was evacuated and refilled with a balloon of H_2 gas (× 3). The resulting reaction mixture was stirred for 3 hours at room temperature then filtered through a plug of celite® to afford a crude, pale yellow gum which was used without further purification.

Step 4: **158** (1.26 g, 2.7 mmol) was dissolved in DMF (27 mL) and **159** (2.22 g, 5.4 mmol) was added and the reaction was stirred at room temperature for 2 hours. A further portion of **159** (0.55 g, 1.4 mmol) was added and the reaction stirred for a further 1 hour then the solvent removed. Purification *via* flash column chromatography on silica gel (0-5% MeOH:CH_2Cl_2) yielded **157** (1.04 g, 1.5 mmol, 20% over 4 steps) as a yellow gum. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 6.96 (s, 1H), 6.83 (s, 1H), 6.80 – 6.73 (m, 2H), 6.10 – 6.02 (m, 2H), 5.17 (s, 1H), 4.53 – 4.48 (m, 1H), 4.31 – 4.28 (m, 2H), 4.23 – 4.18 (m, 4H), 3.61 (app. t, *J* = 5.1 Hz, 4H), 3.40 (s, 6H), 3.15 – 3.10 (m, 1H), 3.08 – 3.02 (m, 2H), 2.93 – 2.88 (m, 1H), 2.71 (d, *J* = 12.9 Hz, 1H), 2.65 (d, *J* = 4.8 Hz, 2H), 2.56 (t, *J* = 6.7 Hz, 2H), 2.21 (s, 3H), 2.13 (t, *J* = 7.4 Hz, 2H), 2.10 (s, 6H), 1.77 – 1.71 (m, 2H), 1.68 – 1.56 (m, 4H), 1.44 – 1.36 (m, 2H); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 172.9, 168.1, 163.7, 147.7, 138.5, 137.3, 135.3, 131.9, 129.7, 124.5, 100.2, 70.9, 62.7, 62.0, 60.3, 59.1, 55.7, 51.0, 40.7, 38.6, 38.0, 36.2, 34.9, 32.8, 31.5, 28.3, 25.7, 19.6, 18.9 ; **HRMS**: $(ESI)^+$ (m/z) Found (M+H): 701.3577, Calc'd $C_{36}H_{53}O_8N_4S^+$ requires 701.3579; **IR** ν_{max}/cm^{-1} (film): 3300, 2933, 1691, 1642, 1490, 1211.

5.5.3.2 Experimental procedures for photoredox-mediated C-benzylation

Photochemical benzylation procedure of exenatide conjugate **167**



A 5 mL microwave tube was charged with a solution of aqueous, unpurified conjugate **86** (400 μM, 50 μL). To the resulting solution was added a magnetic stirrer and Hantzsch ester **155** (2.5 mg, 6.0×10^{-3} mmol, final sol'n conc = 30 mM). The tube was then capped with a septum and subsequently evacuated and refilled with N_2 (× 5). To the tube was added added $[Ru(bpy)_3]^{2+}Cl_2 \bullet 6H_2O$ (5 mM in 3:1 CH_3CN:H_2O, freshly degassed with N_2, 40 μL, final conc = 1 mM), freshly degassed CH_3CN (70 μl), and freshly degassed H_2O (20 μL) . To this resulting mixture was added $K_2S_2O_8$ (100 mM in H_2O, 20 μL, final conc = 10 mM), and the mixture was immediately irradiated with a Kessil 40 W light bulb that was placed at a 5 cm distance from the microwave tube for 5 minutes. A constant temperature was maintained by placing the microwave tube under a stream of air over the course of the irradiation. The reaction mixture was then extracted with 1:1 ethyl acetate:diethyl ether (2 × 0.5 mL) and the excess volatiles were removed *in vacuo*. The resulting solution was then analysed directly *via* LC/MS, (C4 column, Gradient: 5-100% B over 5.2 min then hold for 1 min, then 100-5%B over 1 min, 0.2 ml/min). The reaction was repeated 5 times by two different contributors and was found to have converted to alkylated conjugate **167** in an average of 86% (range from 82-90%) over the 5 runs.

(Isolation of **167** was carried out by Dr M. T. Taylor).[254]

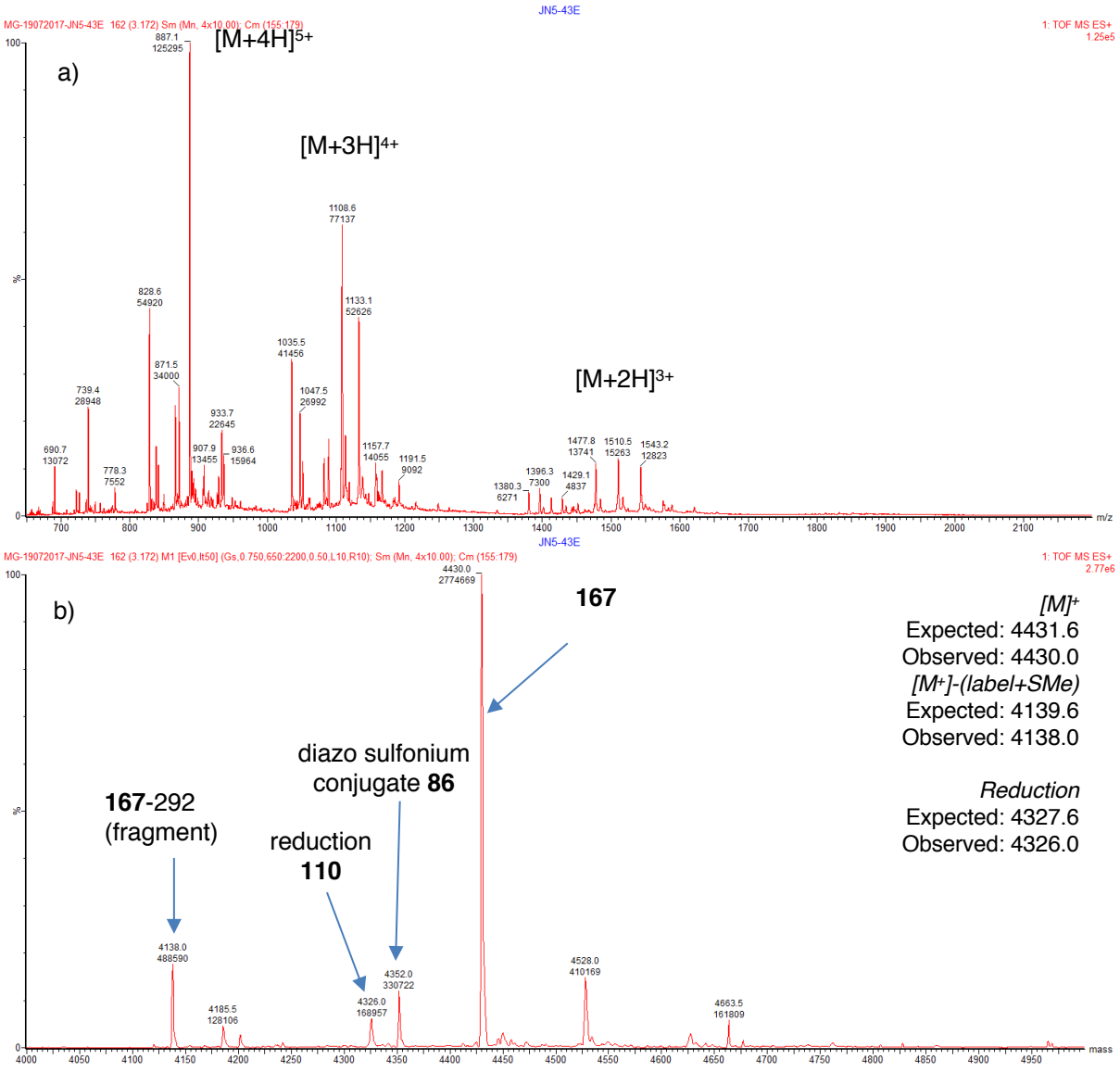
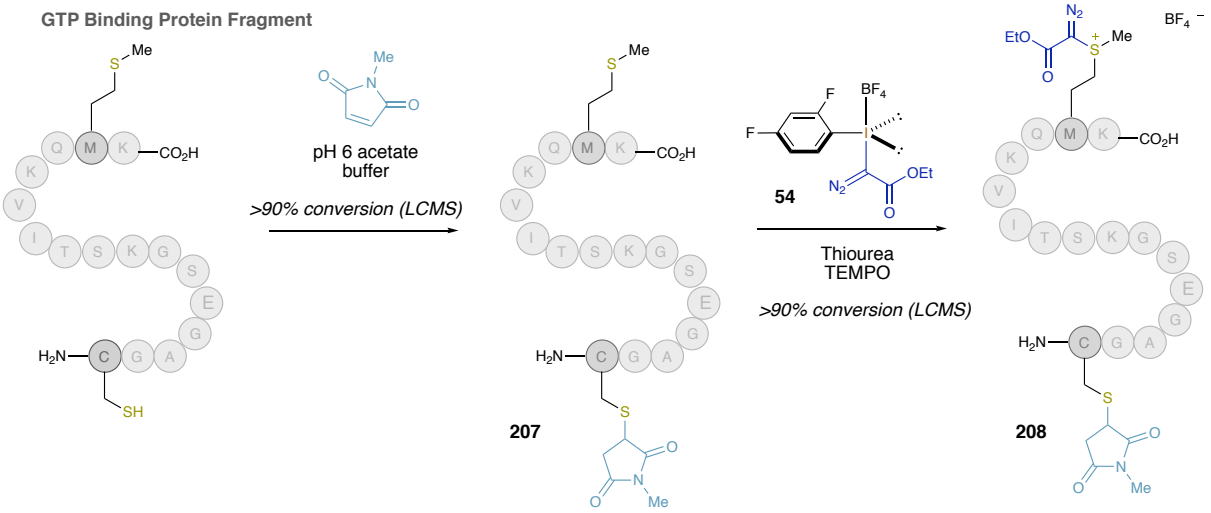


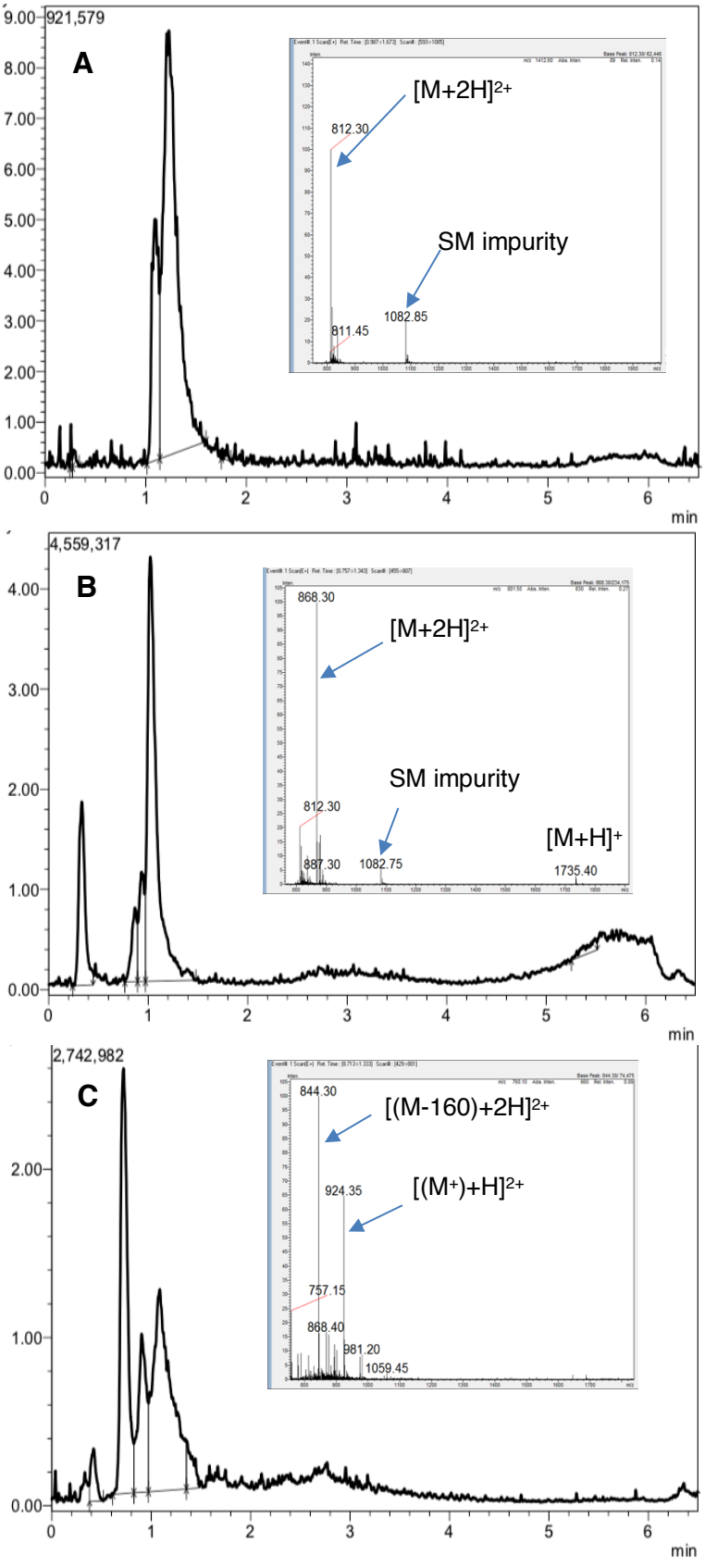
Figure 59 a) Mass chromatogram of the crude reaction of exenatide conjugate **86** in C-benzylation reaction. b) Deconvoluted mass chromatogram.

5.6 Experimental procedures for the multi-functionalisation of polypeptides

5.6.1 Dual labelling of methionine and cysteine



A 2 mL vial equipped with a magnetic stirrer was charged with a solution of GTP Binding Protein Fragment Gα (1 mM in pH 6 buffer, 50 μL). To the vial was added a solution of N-methyl maleimide (10 mM in MeCN, 50 μL) and the resulting solution stirred vigorously for 30 minutes at room temperature, then analysed directly *via* LCMS C18 column ((150 × 10 mm, 10 μm), gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 mL/min), and was observed to have proceeded in >90% conversion in relation to starting material (Figure 60b). A 2 mL vial equipped with a magnetic stirrer was charged with the crude reaction solution (500 μM in 1:1 pH6 buffer:MeCN, 100 μL). To the vial was added thiourea (100 mM in H_2O, 30 μL) and TEMPO (67 mM in H_2O, 30 μL), immediately followed by iodonium salt **54** (250 mM in H_2O, 40 μL) and the reaction stirred vigorously for 2 minutes at room temperature. The resulting mixture was extracted with ethyl acetate:diethyl ether (1:1) (2 × 0.5 mL). The remaining organic volatiles were removed *in vacuo*, and the aqueous layer analysed directly *via* LC/MS C18 column ((150 × 10 mm, 10 μm), gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 mL/min), and was observed to have proceeded in >90% conversion in relation to starting material (Figure 60c).



[M]+
Expected: 1623.9
Observed: 1622.6

[M]+
Expected: 1734.9
Observed: 1735.4

[M]+
Expected: 1847.9
Observed: 1847.7
[M+]-(label+SMe)
Expected: 1687.9
Observed: 1687.6

Figure 60 A) GTP BPF Gα starting material (purity ~75%). B) Crude mass spectrum of step I, reaction with maleimide to form *int*-1. C) Crude mass spectrum of reaction of **207** with **54**. All with mass spectrum shown as an insert.

5.6.2 Dual labelling of methionine and tryptophan

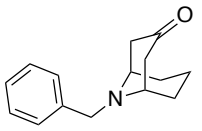
***General Procedure I**: simultaneous dual-labelling of methionine and tryptophan*

A 2 mL vial equipped with a magnetic stirrer was charged with a solution containing the desired polypeptide/protein (2 mM in 0.1M thiourea (aq), 20 μL). To the solution was added *keto*ABNO (40 mM in H_2O, 50 μL) and H_2O (10 μL) and the vial was stirred vigorously before the iodonium salt (100 mM in H_2O, 100 μL) was added and the resulting solution stirred for 30 minutes at room temperature. Ascorbic acid was added (5 mM, 50 μL) and the resulting mixture was then extracted with diethyl ether:ethyl acetate (1:1). The volatile organics were then removed *in vacuo*. The resulting solution was analysed directly *via* LC/MS.

***General Procedure J**: simultaneous dual-labelling of methionine and tryptophan*

A 2 mL vial equipped with a magnetic stirrer was charged with a solution containing the desired polypeptide/protein (2 mM in 0.1M thiourea (aq), 20 μL). To the solution was added *keto*ABNO (40 mM in H_2O, 50 μL) and formic acid (10 μL) before the iodonium salt (100 mM in H_2O, 100 μL) was added and the resulting solution stirred for 30 minutes at room temperature. Ascorbic acid was added (5 mM, 50 μL) and the resulting mixture was then extracted with diethyl ether:ethyl acetate (1:1). The volatile organics were then removed *in vacuo*. The resulting solution was analysed directly *via* LC/MS.

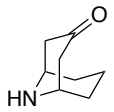
(1R,5S)-9-Benzyl-9-azabicyclo[3.3.1]nonan-3-one **200**



A solution of benzylamine hydrochloride (5.65 g, 39.5 mmol) and glutaraldehyde (13.2 mL, 32.9 mmol, 25% in H_2O) in H_2O (15 mL) was cooled to 0 °C, under air. Acetone dicarboxylic acid (4.80 g, 32.9 mmol) and NaOAc (11.3 mL, 10% aq.) were added and the reaction mixture was stirred for 2 hours at room temperature then heated to 50 °C and stirred overnight. $CHCl_3$ (50 mL) was added and the organic layer extracted. The aqueous layer was extracted with further portions of $CHCl_3$ (3 × 50 mL). The combined organic layers were washed with brine (75 mL), dried over $MgSO_4$, filtered and evaporated to yield the crude product. Purification via flash column chromatography on silica gel (30 – 50% EtOAc/40-60 petroleum ether) afforded the title compound (2.14 g, 28%) as a white crystalline solid. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 7.45 – 7.41 (m, 2H), 7.39 – 7.33 (m, 2H), 7.32 – 7.27 (m, 1H), 3.93 (s, 2H), 3.36 – 3.31 (m, 2H), 2.76 (dd, J = 16.6, 6.7 Hz, 2H), 2.28 (d, J = 16.6 Hz, 2H), 2.04 – 1.91 (m, 2H), 1.65 – 1.52 (m, 4H); **^{13}C-NMR** (100 MHz, $CDCl_3$) δ: 211.8, 139.4, 128.5, 128.4, 127.3, 57.2, 53.7, 43.0, 29.5, 16.8;

*Data in accordance with literature[281]

(1R,5S)-9-Azabicyclo[3.3.1]nonan-3-one **201**

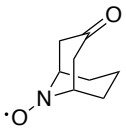


Palladium on carbon (0.43 g, 0.4 mmol, 10 mol% Pd) was added to a solution of **200** (1.84 g, 8.0 mmol) in MeOH (80 mL) under N_2. The flask was evacuated and refilled with a balloon of H_2 gas (× 3). The resulting reaction mixture was stirred for 3 hours at 50 °C then cooled to room temperature and filtered through a plug of celite® to afford the title compound (1.02 g, 93%) as a pale yellow gum which formed a semi-solid on standing. **^{1}H-NMR** (400 MHz, $CDCl_3$) δ: 3.63 – 3.57 (m, 2H), 2.61 (dd, J = 16.5, 6.8 Hz, 2H), 2.41 (d, J = 16.5 Hz, 2H), 1.83 – 1.74 (m, 3H), 1.81 – 1.77 (m, 2H), 1.70 – 1.62 (m, 2H), 1.59 – 1.53 (m, 2H); **^{13}C-NMR** (100 MHz, $CDCl_3$)

δ: 211.1, 49.3, 47.5, 32.0, 16.7; **HRMS**: $(ESI)^+$ (m/z) $[M+2H]^+$: Found 140.1078, Calc. $C_8H_{14}NO^+$ requires 140.1075;

*Data in accordance with literature[281]

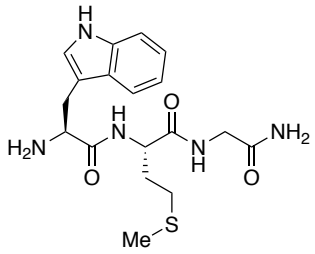
9-Azabicyclo[3.3.1]nonan-3-one N-oxyl (keto-ABNO) **185**



To a solution of **201** (0.70 g, 5.0 mmol) in acetonitrile (17 mL), $Na_2WO_4 \cdot 2H_2O$ (0.17 g, 0.5 mmol) and urea hydrogen peroxide (3.68 g, 39.0 mmol) were added and the reaction mixture was stirred at room temperature for 6 hours. H_2O was added to the reaction mixture and the aqueous solution was extracted with $CHCl_3$ (3 × 25 mL). The organic layer was dried over K_2CO_3 and concentrated. The residue was purified by flash column chromatography on silica gel (50% EtOAc/46-60 petroleum ether) followed by recrystallisation from EtOAc/40-60 petroleum ether to afford the title compound (0.52 g, 68%) as a yellow solid. **HRMS**: $(ESI)^+$ (m/z) $[M+2H]^+$: Found 154.0866, Calc. $C_8H_{12}NO_2^+$ requires 154.0868; **IR**: 2964, 1704, 1353, 1192;

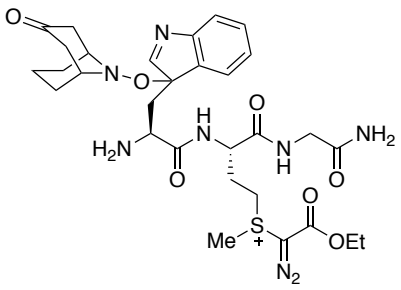
*Data in accordance with literature[281]

Tripeptide **191**



Peptide synthesis was carried out according to literature procedures,[282] using Rink amide MBHA resin (0.40 – 0.80 mmol g^{-1}) for Fmoc-based solid-phase synthesis with 20% piperidine/DMF deprotection and HCTU/DIPEA activation. The resulting resin-bound peptide was cleaved and side-chains deprotected using a solution of TFA, triisopropylsilane, CH_2Cl_2, H_2O (92.5:2.5:2.5:2.5) for 2 hours. The resin was filtered from the cleaved peptide and the filtrate concentrated by bubbling N_2 gas through the solution. The crude peptide was precipitated with diethyl ether, centrifuged, dried in vacuo and purified using reverse phase semi-preparative HPLC (C18 column (150 × 10 mm, 10µm), Gradient: 5-10% B over 9 min then 10-95% B over 2 min, hold for 1 min, then 95-5% B over 2 min, hold for 1 min, 5 ml/min), followed by lyophilisation to yield the title compound as a white solid (0.12 g, 52%). **^{1}H-NMR** (500 MHz, CD_3CN) δ: 9.43 (br. s, 1H), 7.88 (d, J = 7.1 Hz, 1H), 7.63 (d, J = 7.6 Hz, 1H), 7.43 – 7.40 (m, 1H), 7.35 – 7.30 (m, 1H), 7.23 (d, J = 2.4 Hz, 1H), 7.17 – 7.12 (m, 1H), 7.09 – 7.05 (m, 1H), 6.56 (br. s, 1H), 5.94 (br. s, 1H), 4.38 – 4.31 (m, 1H), 4.27 (t, J = 7.1 Hz, 1H), 3.77 (dd, J = 16.9, 6.1 Hz, 1H), 3.66 (dd, J = 16.9, 5.7 Hz, 1H), 3.39 (dd, J = 14.8, 6.3 Hz, 1H), 32.6 (dd, J = 14.8, 7.6 Hz, 1H), 2.51 – 2.33 (m, 4H), 2.02 (s, 3H); **^{13}C-NMR** (125 MHz, $CDCl_3$) δ: 172.8, 172.2, 162.8, 137.7, 128.3, 125.7, 122.9, 120.2, 119.5, 112.5, 109.5, 55.6, 54.0, 43.2, 31.9, 30.6, 29.4, 15.3; **HRMS**: $(ESI)^+$ (m/z) $[M+H]^+$: Found 392.1745, Calc. $C_{18}H_{26}N_5O_3S^+$ requires 392.1751; **IR**: 3299, 1667, 1529, 1201.

Dual labelled tripeptide **192**/**192'**



A 2 mL vial equipped with a magnetic stirrer was charged with a solution containing the tripeptide (0.2 mM in H_2O, 25 µL). To the solution was added ketoABNO (40 mM in H_2O, 500 µL), thiourea (0.1 M in H_2O, 400 µL) and H_2O (75 µL) and the vial was stirred vigorously, before iodonium salt **54** (100 mM in CH_3CN, 1.0 mL) was added and the resulting solution stirred for 15 minutes at room temperature. The reaction mixture was then extracted with diethyl ether:ethyl acetate (1:1) (2 × 0.5 mL). The volatile organics were then removed in vacuo. The resulting solution was analysed directly via LC/MS C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min) and was observed to have proceeded in >95% conversion. Purification via reverse-phase HPLC (C18 column (150 × 10 mm, 10µm), Gradient: 5-95% B over 11 min then hold for 0.5 min, then 95-5%B over 0.5 min, hold for 3 min, 5 ml/min), followed by lyophilization yielded **192** as an off-white solid that was immediately redissolved in H_2O (500 µL) and stored at –20°C. ***192** was isolated as a mixture of 3 products; one hydrated and two dehydrated forms, which made NMR analysis intractable. **HRMS**: $(ESI)^+$ (m/z) $[M]^+$: Found 657.2819, Calc. $C_{30}H_{41}N_8O_7S^+$ requires 657.2813;

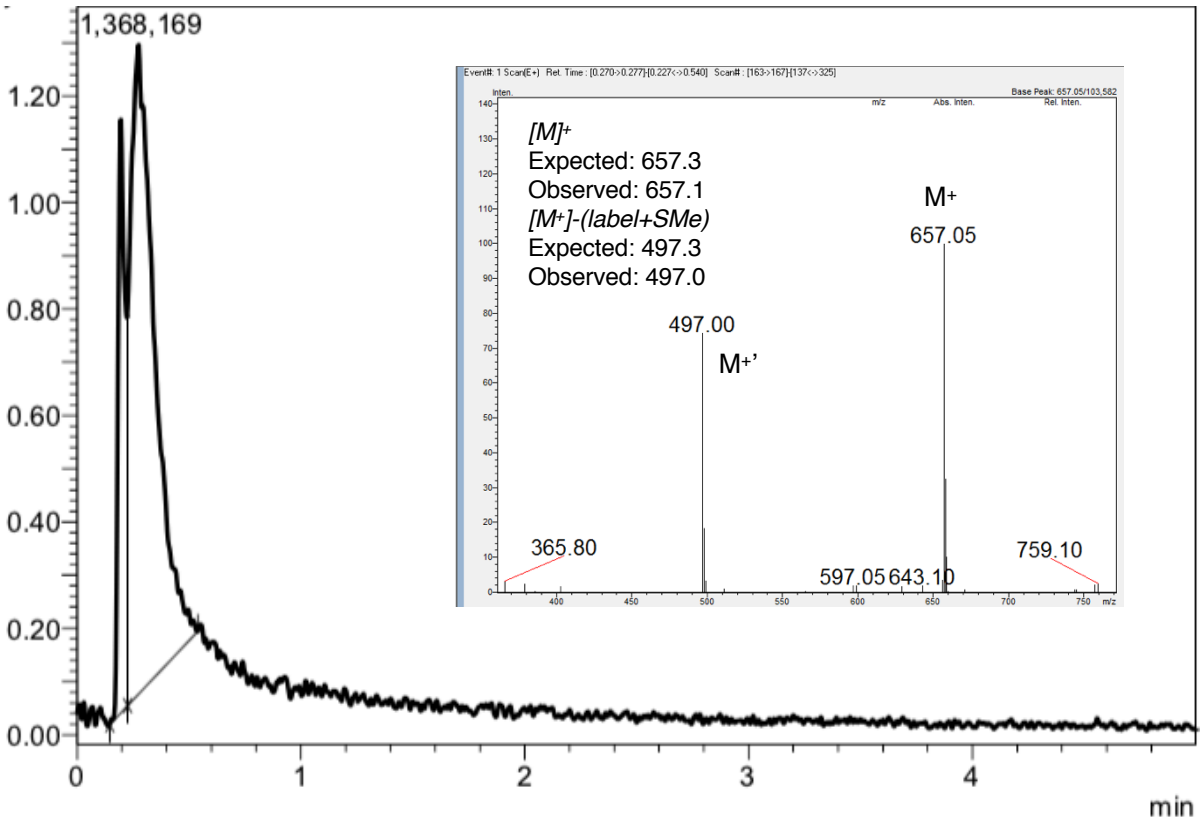
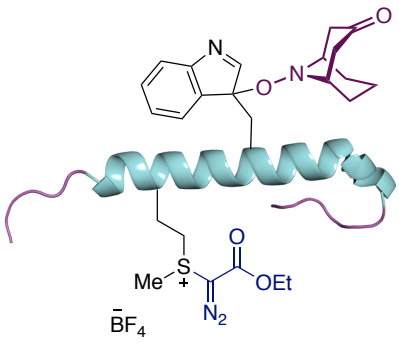


Figure 61 Mass trace of crude reaction of tripeptide with **54**, with mass chromatogram shown as an insert. ' denotes diagnostic fragmentation of C-S bond giving M+label-(label+SMe) = M+113-160.

5.6.3 Substrate scope for the multi-functionalisation of polypeptides

Dual labelling of exenatide to form **193**, **194** & **194'**



The dual labelling of exenatide was performed at room temperature using General Procedure I using **54** (100 mM in H_2O) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) in the extraction step. The resulting solution was then analysed directly *via* LC/MS C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min) and was observed to have proceeded in 83-92% conversion (n = 5).

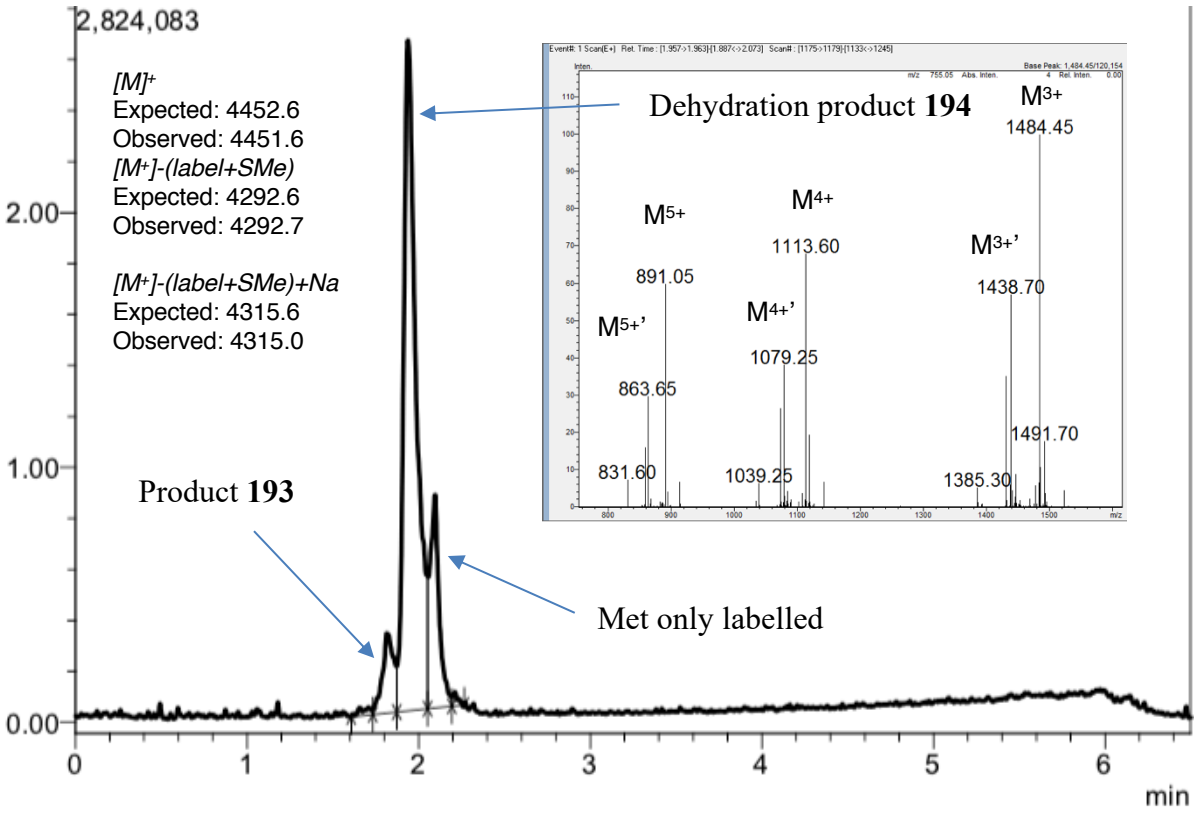
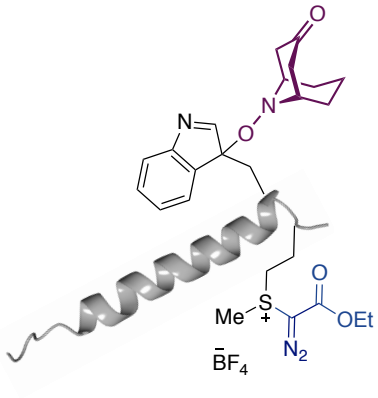


Figure 62 Mass trace of crude reaction of exenatide with **54**, with mass chromatogram shown as an insert. ' denotes diagnostic fragmentation of C-S bond giving M+label-(label+SMe) = M+113-160.

Dual labelling of glucagon to form **202**



The dual labelling of glucagon was performed at room temperature using General Procedure J using **54** (100 mM in H_2O) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) in the extraction step. The resulting solution was then analysed directly via LC/MS C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min) and was observed to have proceeded in 77% conversion.

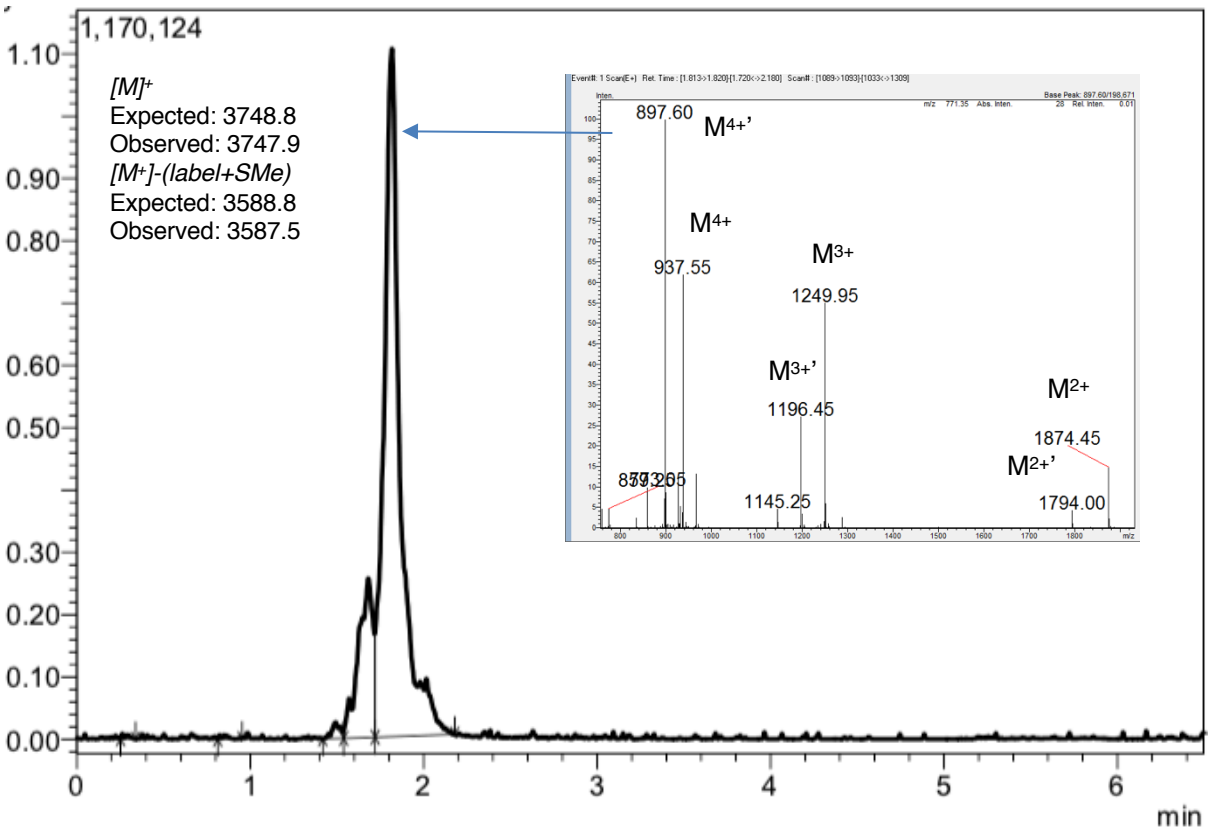
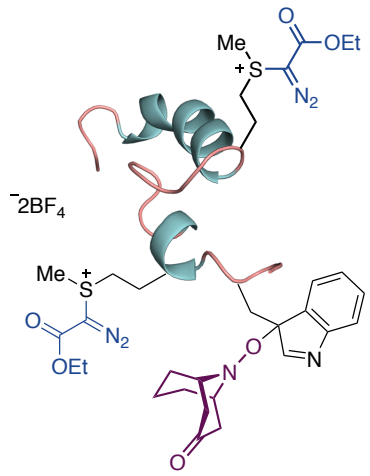


Figure 63 Mass trace of crude reaction of glucagon with **54**, with mass chromatogram shown as an insert. ' denotes diagnostic fragmentation of C-S bond giving M+label-(label+SMe) = M+113-160.

Dual labelling of teriparatide to form **203**



The dual labelling of teriparatide was performed at room temperature using General Procedure J using **54** (100 mM in H_2O) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) in the extraction step. The resulting solution was then analysed directly via LC/MS C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min) and was observed to have proceeded in >95% conversion.

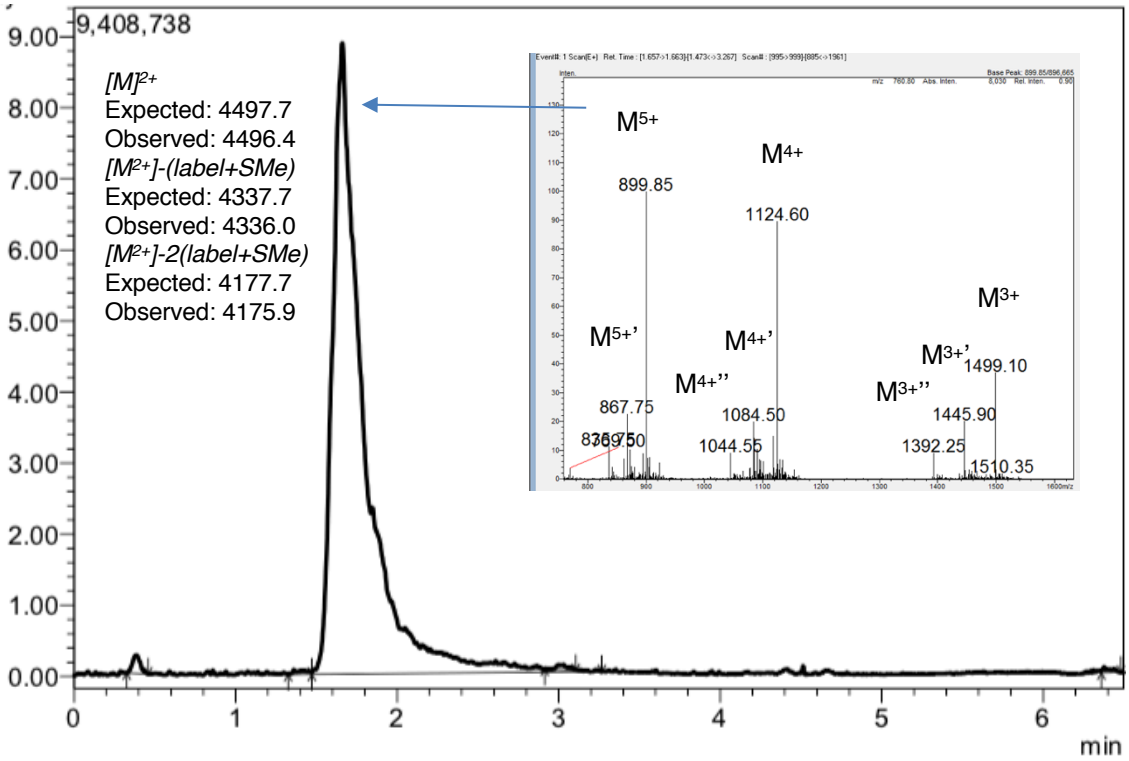
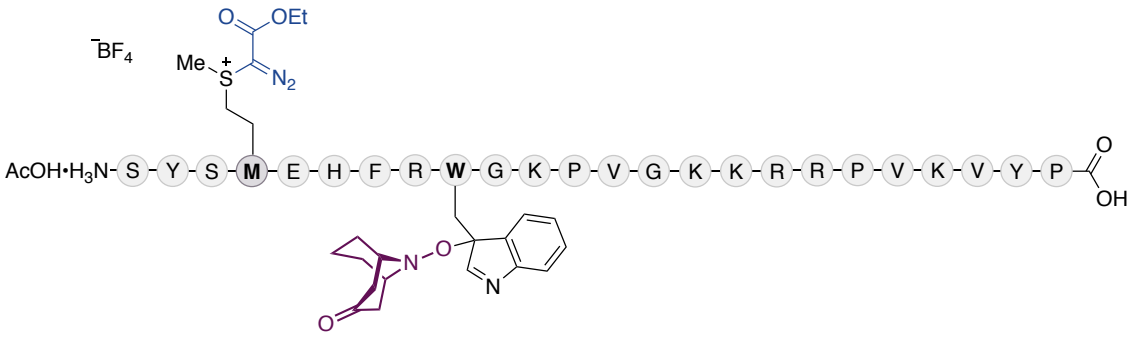


Figure 64 Mass trace of crude reaction of teriparatide with **54**, with mass chromatogram shown as an insert. ' denotes diagnostic fragmentation of C-S bond giving M+label-(label+SMe) = M+(154+2(113))-160. '' denotes diagnostic fragment M+(154+2(113))-2(160).

Dual labelling of tetracosactide to form **204**



The dual labelling of tetracosactide was performed at room temperature using General Procedure I using **54** (100 mM in H_2O) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) in the extraction step. The resulting solution was then analysed directly via LC/MS C18 column (50 × 2.1 mm, 2.6 µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min) and was observed to have proceeded in 63% conversion.

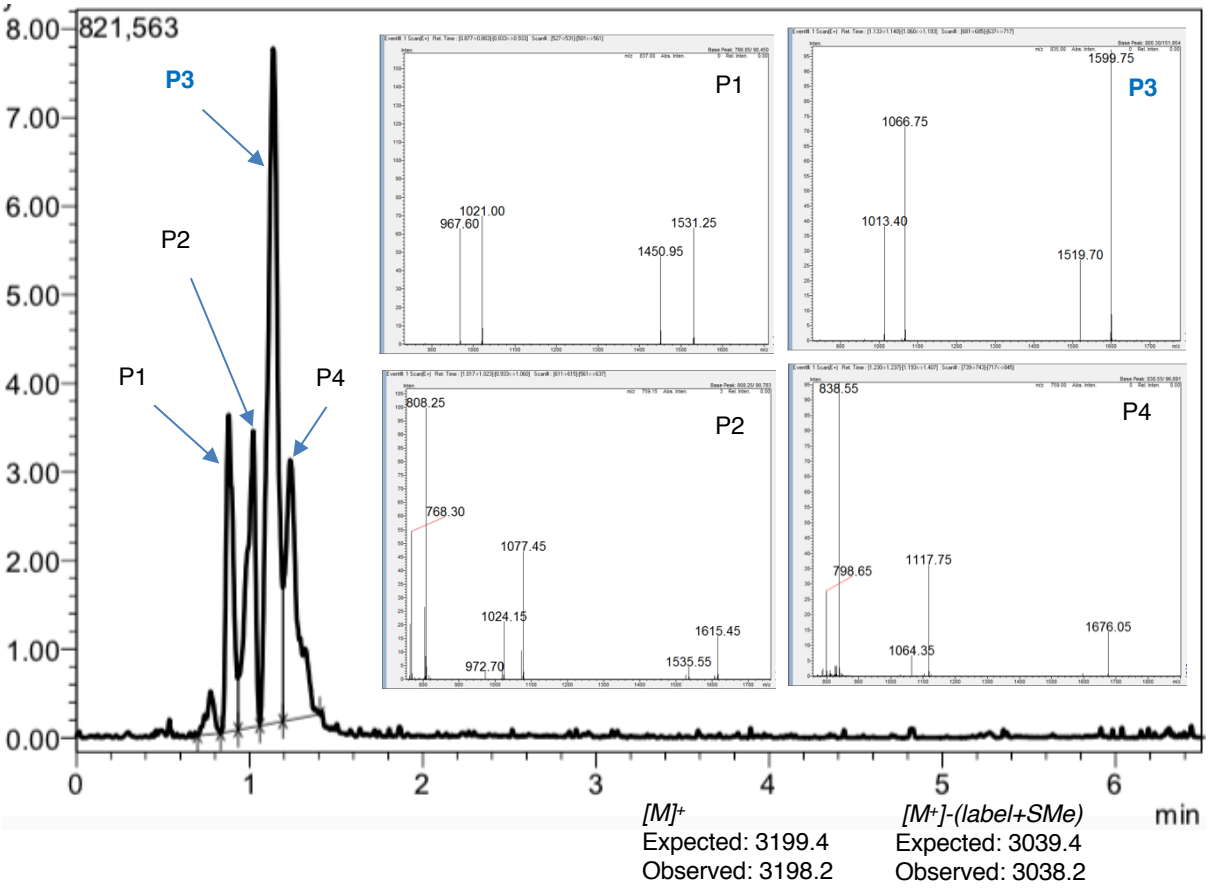
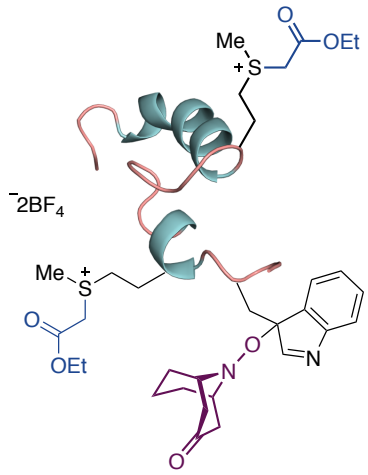


Figure 65 Mass trace of crude reaction of tetracosactide with **54**, with mass chromatograms shown as an insert. P1 = tetracosactide +met label+O (14% conversion). P2 = tetracosactide+Met label+Trp label+29 (10% conversion). **P3** = desired product **204** (tetracosactide+Met label+Trp label) (63% conversion). P4 = tetracosactide+Met label+ 2(Trp label) (12% conversion).

Preliminary result for photoreduction of multi-functionalised teriparatide conjugate **203** *to form* **205**



The dual labelling of teriparatide was performed at room temperature using General Procedure J using **54** (100 mM in H_2O) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) in the extraction step. To the resulting solution was added a magnetic stirrer and the Hantzsch ester **102** (1.8 mg, 1.0×10^{-5} mmol). The tube was then capped with a septum and subsequently evacuated/backfilled with N_2 (× 5). To the tube was added added fac-Ir(ppy)$_3$ (560 µM in freshly degassed CH_3CN, 160 µL) and freshly degassed H_2O (160 µL). The resulting heterogeneous mixture was then stirred rapidly. The mixture was irradiated with a Prolite 30 W fluorescent light bulb that was placed at a 5 cm distance from the microwave tube for 45 minutes. A constant temperature was maintained by placing the microwave tube under a stream of air over the course of the irradiation. The reaction mixture was then extracted with 1:1 ethyl acetate:diethyl ether (3 × 0.5 mL) and the excess volatiles were removed in vacuo. The resulting solution was then analysed directly via LC/MS (C18 column (50 × 2.1 mm, 2.6µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min), and the doubly reduced species was observed as the major product.

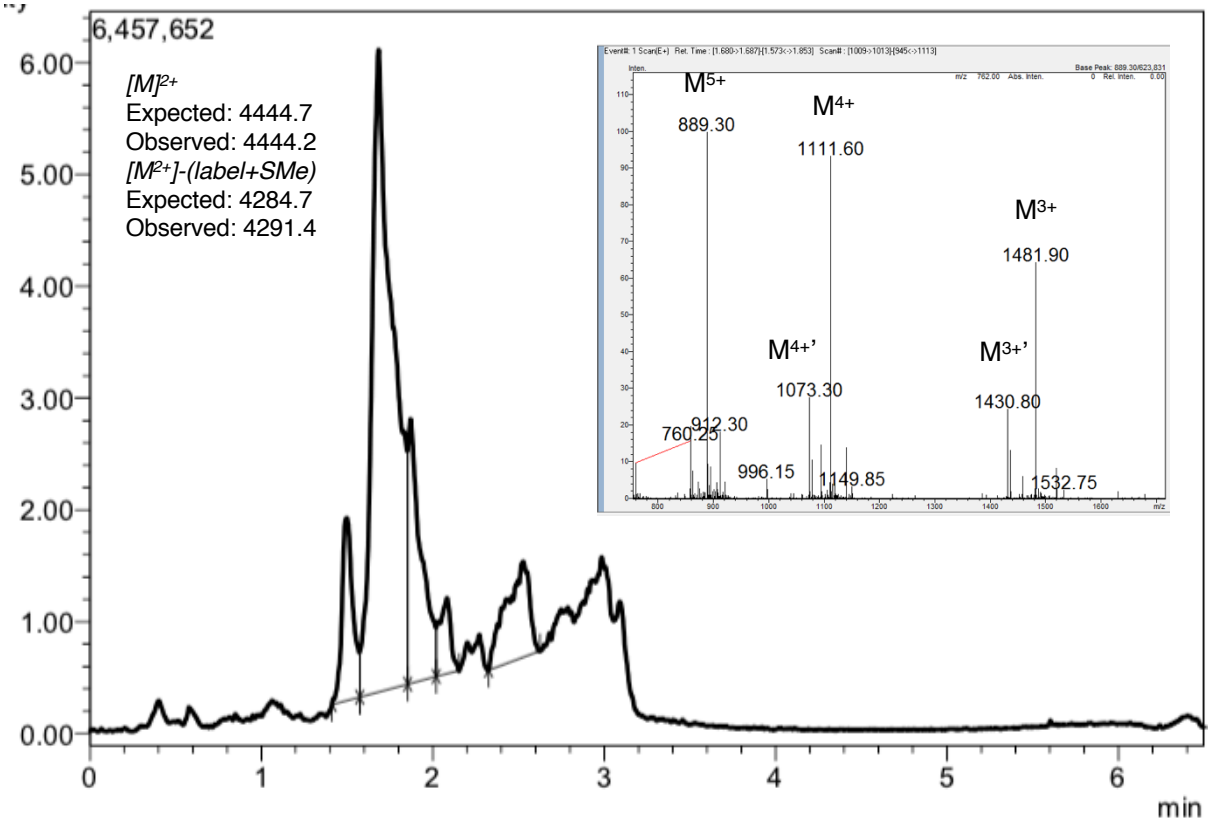
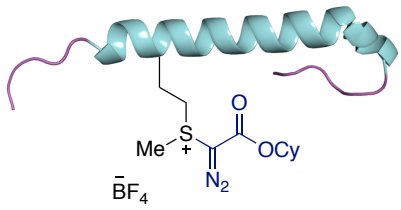


Figure 66 Mass trace of crude reaction of teriparatide with **54** followed by photoreduction, with mass chromatogram shown as an insert. ' denotes diagnostic fragmentation of C-S bond giving M+label-(label+SMe) = M+(154+2(113))-160. '' denotes diagnostic fragment M+(154+2(113))-2(160).

Photoreduction of exenatide conjugate **86** *with photocatalyst-free conditions*



The labelling of exenatide was performed at room temperature using General Procedure E using **77** (100 mM in H_2O) and 1:1 diethyl ether:ethyl acetate (2 × 0.5 mL) in the extraction step. To the resulting solution was added a magnetic stirrer and the Hantzsch ester **102** (1.0 mg, 4.0 × 10^{-6} mmol). The tube was then capped with a septum and subsequently evacuated/backfilled with N_2 (× 5). To the tube was added freshly degassed CH_3CN (100 µL) and freshly degassed H_2O (50 µL). The resulting heterogeneous mixture was then stirred rapidly. The mixture was irradiated with a Kessil 40 W light bulb that was placed at a 5 cm distance from the microwave tube for 5 minutes. A constant temperature was maintained by placing the microwave tube

under a stream of air over the course of the irradiation. The reaction mixture was then extracted with 1:1 ethyl acetate:diethyl ether (3 × 0.5 mL) and the excess volatiles were removed *in vacuo*. The resulting solution was then analysed directly *via* LC/MS (C18 column (50 × 2.1 mm, 2.6µm), Gradient: 5-95% B over 5 min then hold for 0.5 min, 0.7 ml/min), and reaction was observed to have proceeded in >95% conversion.

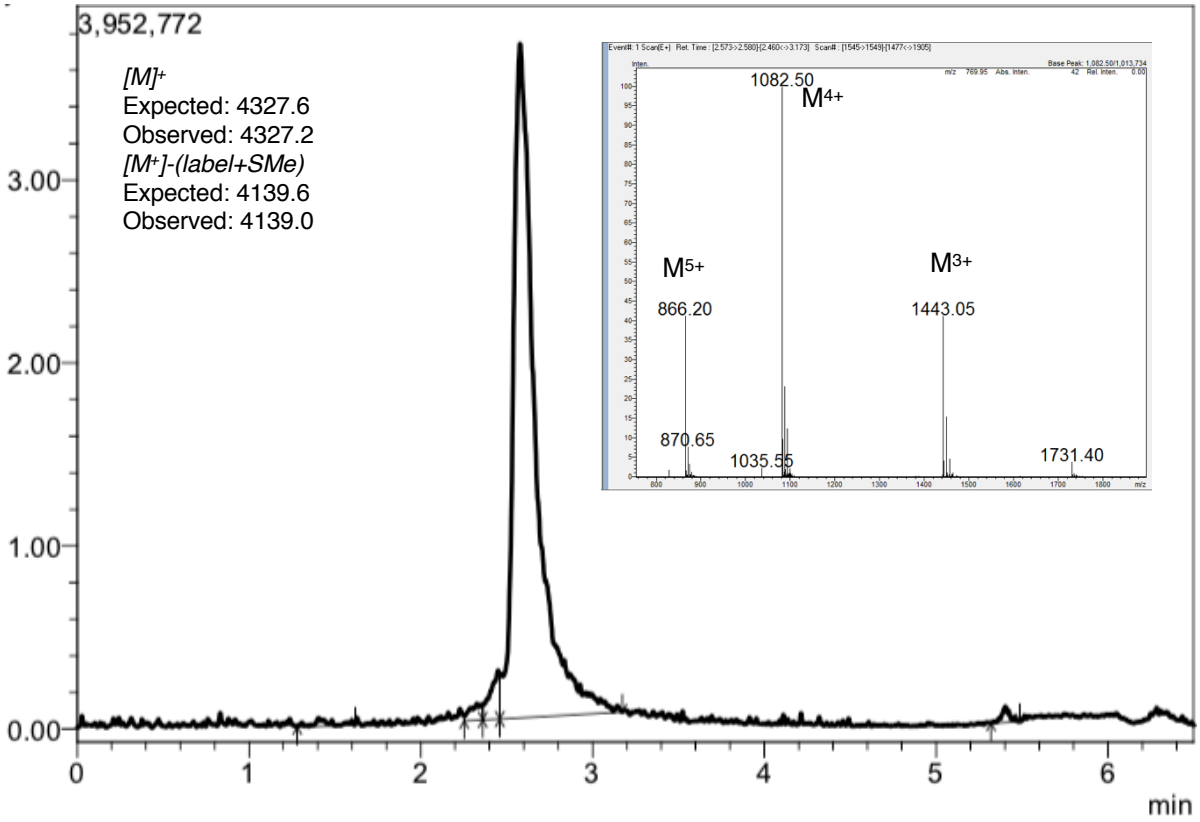


Figure 67 Mass trace of crude reaction of exenatide with **54**, followed by reduction, with mass chromatogram shown as an insert.

6 References

(1) McKay, C. S.; Finn, M. G. *Chem. Biol.* **2014**, *21* (9), 1075–1101.

(2) Koniev, O.; Wagner, A. *Chem. Soc. Rev.* **2015**, *44* (15), 5495–5551.

(3) Stephanopoulos, N.; Francis, M. B. *Nat. Chem. Biol.* **2011**, *7* (12), 876–884.

(4) Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2009**, *48* (38), 6974–6998.

(5) Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. *Chem. Asian J.* **2009**, *4* (5), 630–640.

(6) Spicer, C. D.; Davis, B. G. *Nat. Commun.* **2014**, *5*, 1–14.

(7) Carrico, I. S. *Chem. Soc. Rev.* **2008**, *37* (7), 1423–1431.

(8) Lang, K.; Chin, J. W. *ACS Chem. Biol.* **2014**, *9* (1), 16–20.

(9) Lang, K.; Chin, J. W. *Chem. Rev.* **2014**, *114* (9), 4764–4806.

(10) Wang, L.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2005**, *44* (1), 34–66.

(11) Chin, J. W. *Nature* **2017**, *550* (7674), 53–60.

(12) Lang, K.; Chin, J. W. *Chem. Commun.* **2010**, *46* (10), 1589–1600.

(13) Sletten, E. M.; Bertozzi, C. R. *Acc. Chem. Res.* **2011**, *44*, 666–676.

(14) Li, F.; Mahato, R. I. *Mol. Pharm.* **2017**, *14*, 1321–1324.

(15) Uttamchandani, M.; Li, J.; Sun, H.; Yao, S. Q. *ChemBioChem* **2008**, *9* (5), 667–675.

(16) Speers, A. E.; Cravatt, B. F. *ChemBioChem* **2004**, *5* (1), 41–47.

(17) Prescher, J. A.; Dube, D. H.; Bertozzi, C. R. *Nature* **2004**, *430*, 873–877.

(18) King, M.; Wagner, A. *Bioconjugate Chem.* **2014**, *25*, 825–839.

(19) Devaraj, N. K. *ACS Cent. Sci.* **2018**, *4*, 952–959.

(20) Hermanson, G. T. *Bioconjugate Techniques*, 2nd edition.; Academic Press, 2008.

(21) Brotzel, F.; Mayr, H. *Org. Biomol. Chem.* **2007**, *5* (23), 3814–3820.

(22) Goddard, D. R.; Michaelis, L. *J. Biol. Chem.* **1935**, *112* (6), 361–371.

(23) King, J. L.; Jukes, T. H. *Science* **1969**, *164* (3881), 788–798.

(24) Dan, N.; Setua, S.; Kashyap, V. K.; Khan, S.; Jaggi, M.; Yallapu, M. M.; Chauhan, S. C. *Pharmaceuticals* **2018**, *11* (2), 32–54.

(25) Góngora-Benítez, M.; Tulla-Puche, J.; Albericio, F. *Chem. Rev.* **2014**, *114* (2), 901–926.

(26) Mamathambika, B. S.; Bardwell, J. C. *Annu. Rev. Cell Dev. Bi.* **2008**, *24* (1), 211–235.

(27) Petersen, M. T. N.; Jonson, P. H.; Petersen, S. B. *Protein Eng. Des. Sel.* **1999**, *12* (7), 535–548.

(28) Burns, J. A.; Butler, J. C.; Moran, J.; Whitesides, G. M. *J. Org. Chem.* **1991**, *56* (8), 2648–2650.

(29) Getz, E. B.; Xiao, M.; Chakrabarty, T.; Cooke, R.; Selvin, P. R. *Anal. Biochem.* **1999**, *273* (1), 73–80.

(30) Rosendahl, M. S.; Doherty, D. H.; Smith, D. J.; Carlson, S. J.; Chlipala, E. A.; Cox, G. N. *Bioconjugate Chem.* **2005**, *16*, 200–207.

(31) Glockshuber, R.; Schmidt, T.; Plueckthun, A. *Biochemistry* **1992**, *31* (5), 1270–1279.

(32) Friedmann, E.; Marrian, D. H.; Simon-reuss, I. *Brit. J. Pharmacol.* **1949**, *4* (1), 105–108.

(33) Friedmann, E. *Biochim. Biophys. Acta* **1952**, *9*, 65–75.

(34) Philippe, S.; Christophe, B.; Benoît, F.; Francis, S. *Bioconjugate Chem.* **2000**, *11*, 118–123.

(35) Tedaldi, L. M.; Smith, M. E. B.; Nathani, R. I.; Baker, J. R. *Chem. Commun.* **2009**, *43*, 6583–6585.

(36) Zhang, Y.; Bhatt, V. S.; Sun, G.; Wang, P. G.; Palmer, A. F. *Bioconjugate Chem.* **2008**, *19* (11), 2221–2230.

(37) Saito, F.; Noda, H.; Bode, J. W. *ACS Chem. Biol.* **2015**, *10*, 1026–1033.

(38) Alley, S. C.; Okeley, N. M.; Senter, P. D. *Curr. Opin. Chem. Biol.* **2010**, *14* (4), 529–537.

(39) Kim, J.-R.; Yoon, H. W.; Kwon, K.-S.; Lee, S.-R.; Rhee, S. G. *Anal. Biochem.* **2000**, *283* (2), 214–221.

(40) Davis, N. J.; Flitsch, S. L. *Tetrahedron Lett.* **1991**, *32* (46), 6793–6796.

(41) Lindley, H. *Nature* **1956**, *178* (4534), 647–648.

(42) Raftery, M. A.; Cole, R. D. *J. Biol. Chem.* **1966**, *241* (15), 3457–3461.

(43) Simon, M. D.; Chu, F.; Racki, L. R.; De La Cruz, C. C.; Burlingame, A. L.; Panning, B.; Narlikar, G. J.; Shokat, K. M. *Cell* **2007**, *128*, 1003–1012.

(44) Smith, H. B.; Hartman, F. C. *J. Biol. Chem.* **1988**, *263* (10), 4921–4925.

(45) Morpurgo, M.; Veronese, F. M.; Kachensky, D.; Milton Harris, J. M. *Bioconjugate Chem.* **1996**, *7* (3), 363–368.

(46) Morales-Sanfrutos, J.; Lopez-Jaramillo, F. J.; Hernandez-Mateo, F.; Santoyo-Gonzalez, F. *J. Org. Chem.* **2010**, *75* (12), 4039–4047.

(47) Ovaa, H.; van Swieten, P. F.; Kessler, B. M.; Leeuwenburgh, M. A.; Fiebiger, E.; van den Nieuwendijk, A. M. C. H.; Galardy, P. J.; van der Marel, G. A.; Ploegh, H. L.; Overkleeft, H. S. *Angew. Chem. Int. Ed.* **2003**, *42* (31), 3626–3629.

(48) Bernardim, B.; Cal, P. M. S. D.; Matos, M. J.; Oliveira, B. L.; Martínez-Sáez, N.; Albuquerque, I. S.; Perkins, E.; Corzana, F.; Burtoloso, A. C. B.; Jiménez-Osés, G.; Bernardes, G. J. L. *Nat. Commun.* **2016**, *7* (13128), 1–9.

(49) Hoyle, C. E.; Bowman, C. N. *Angew. Chem. Int. Ed.* **2010**, *49* (9), 1540–1573.

(50) Dondoni, A.; Marra, A. *Chem. Soc. Rev.* **2012**, *41* (2), 573–586.

(51) Dondoni, A.; Massi, A.; Nanni, P.; Roda, A. *Chem. Eur. J.* **2009**, *15* (43), 11444–11449.

(52) Stolz, R. M.; Northrop, B. H. *J. Org. Chem.* **2013**, *78* (16), 8105–8116.

(53) Massi, A.; Nanni, D. *Org. Biomol. Chem.* **2012**, *10* (19), 3791–3807.

(54) Bader, H.; Cross, L. C.; Heilbron, I.; Jones, E. R. H. *J. Chem. Soc.* **1949**, *0* (0), 619–623.

(55) Lo Conte, M.; Pacifico, S.; Chambery, A.; Marra, A.; Dondoni, A. *J. Org. Chem.* **2010**, *75* (13), 4644–4647.

(56) Conte, M. Lo; Staderini, S.; Marra, A.; Sanchez-Navarro, M.; Davis, B. G.; Dondoni, A. *Chem. Commun.* **2011**, *47* (39), 11086–11088.

(57) Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, *82* (1), 70–77.

(58) Fontana, A.; Scoffone, E.; Benassi, C. A. *Biochemistry* **1968**, *7* (3), 980–986.

(59) Davis, B. G.; Maughan, M. A. T.; Green, M. P.; Ullman, A.; Jones, B. *Tetrahedron-Asymmetr* **2000**, *11*, 245–262.

(60) Gamblin, D. P.; Garnier, P.; Ward, S. J.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. *Org. Biomol. Chem.* **2003**, *1* (21), 3642–3644.

(61) Pollack, S. J.; Schultz, P. G. *J. Am. Chem. Soc.* **1989**, *111* (5), 1929–1931.

(62) Gamblin, D. P.; Garnier, P.; van Kasteren, S.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. *Angew. Chem. Int. Ed.* **2004**, *43* (7), 828–833.

(63) Smith, M. E. B.; Schumacher, F. F.; Ryan, C. P.; Tedaldi, L. M.; Papaioannou, D.; Waksman, G.; Caddick, S.; Baker, J. R. *J. Am. Chem. Soc.* **2010**, *132* (6), 1960–1965.

(64) Castañeda, L.; Wright, Z. V. F.; Marculescu, C.; Tran, T. M.; Chudasama, V.; Maruani, A.; Hull, E. A.; Nunes, J. P. M.; Fitzmaurice, R. J.; Smith, M. E. B.; Jones, L. H.; Caddick, S.; Baker, J. R. *Tetrahedron Lett.* **2013**, *54* (27), 3493–3495.

(65) Jones, M. W.; Strickland, R. A.; Schumacher, F. F.; Caddick, S.; Baker, J. R.; Gibson, M. I.; Haddleton, D. M. *J. Am. Chem. Soc.* **2012**, *134* (3), 1847–1852.

(66) Forte, N.; Chudasama, V.; Baker, J. R. *Drug Discov. Today Technol.* **2018**, *30*, 11–20.

(67) Walsh, S. J.; Omarjee, S.; Galloway, W. R. J. D.; Kwan, T. T.-L.; Sore, H. F.; Parker, J. S.; Hyvönen, M.; Carroll, J. S.; Spring, D. R. *Chem. Sci.* **2019**, *10*, 694–700.

(68) Toda, N.; Asano, S.; Barbas, C. F. *Angew. Chem. Int. Ed.* **2013**, *52* (48), 12592–12596.

(69) Lyon, R. P.; Setter, J. R.; Bovee, T. D.; Doronina, S. O.; Hunter, J. H.; Anderson, M. E.; Balasubramanian, C. L.; Duniho, S. M.; Leiske, C. I.; Senter, P. D. *Nat. Biotechnol.* **2014**, *32* (10), 1059–1062.

(70) Szijj, P. A.; Bahou, C.; Chudasama, V. *Drug Discov. Today Technol.* **2018**, *30*, 27–34.

(71) Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130* (15), 5052–5053.

(72) Dadová, J.; Galan, S. R.; Davis, B. G. *Curr. Opin. Chem. Biol.* **2018**, *46*, 71–81.

(73) Strumeyer, D. H.; White, W. N.; Koshland, D. E. *P. Natl. A. Sci. USA.* **1963**, *50* (5), 931–935.

(74) Okeley, N. M.; Zhu, Y.; Van der Donk, W. A. *Org. Lett.* **2000**, *2* (23), 3603–3606.

(75) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernández-González,

M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. *Chem. Sci.* **2011**, *2* (9), 1666–1676.

(76) Hashimoto, K.; Sakai, M.; Okuno, T.; Shirahama, H. *Chem. Commun.* **1996**, *0* (10), 1139–1140.

(77) Seebeck, F. P.; Szostak, J. W. *J. Am. Chem. Soc.* **2006**, *128* (22), 7150–7151.

(78) Wang, J.; Schiller, S. M.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2007**, *46* (36), 6849–6851.

(79) You, Y. O.; Levengood, M. R.; Ihnken, L. A. F.; Knowlton, A. K.; van der Donk, W. A. *ACS Chem. Biol.* **2009**, *4* (5), 379–385.

(80) Galonić, D. P.; van der Donk, W. A.; Gin, D. Y. *Chem. Eur. J.* **2003**, *9* (24), 5997–6006.

(81) Tamura, T.; Hamachi, I. *J. Am. Chem. Soc.* **2019**, *141* (7), 2782–2799.

(82) Wright, T. H.; Bower, B. J.; Chalker, J. M.; Bernardes, G. J. L.; Wiewiora, R.; Ng, W.-L.; Raj, R.; Faulkner, S.; Vallée, M. R. J.; Phanumartwiwath, A.; Coleman, O. D.; Thézénas, M.-L.; Khan, M.; Galan, S. R. G.; Lercher, L.; Schombs, M. W.; Gerstberger, S.; Palm-Espling, M. E.; Baldwin, A. J.; Kessler, B. M.; Claridge, T. D. W.; Mohammed, S.; Davis, B. G. *Science* **2016**, *354* (6312), aag1465.

(83) Galan, S. R. G.; Wickens, J. R.; Dadova, J.; Ng, W. L.; Zhang, X.; Simion, R. A.; Quinlan, R.; Pires, E.; Paton, R. S.; Caddick, S.; Chudasama, V.; Davis, B. G. *Nat. Chem. Biol.* **2018**, *14* (10), 955–963.

(84) Antos, J. M.; Francis, M. B. *Curr. Opin. Chem. Biol.* **2006**, *10* (3), 253–262.

(85) Simmons, R. L.; Yu, R. T.; Myers, A. G. *J. Am. Chem. Soc* **2011**, *133*, 15870–15873.

(86) Yang, M.; Li, J.; Chen, P. R. *Chem. Soc. Rev.* **2014**, *43* (18), 6511–6526.

(87) Cheng, G.; Lim, R. K. V.; Li, N.; Lin, Q. *Chem. Commun.* **2013**, *49* (60), 6809–6811.

(88) Vinogradova, E. V.; Zhang, C.; Spokoyny, A. M.; Pentelute, B. L.; Buchwald, S. L. *Nature* **2015**, *526*, 687–691.

(89) Zhao, W.; Lee, H. G.; Buchwald, S. L.; Hooker, J. M. *J. Am. Chem. Soc.* **2017**, *139* (21), 7152–7155.

(90) Kubota, K.; Dai, P.; Pentelute, B. L.; Buchwald, S. L. *J. Am. Chem. Soc.* **2018**, *140* (8),

3128–3133.

(91) Rojas, A. J.; Pentelute, B. L.; Buchwald, S. L. *Org. Lett.* **2017**, *19* (16), 4263–4266.

(92) Willwacher, J.; Raj, R.; Mohammed, S.; Davis, B. G. *J. Am. Chem. Soc.* **2016**, *138* (28), 8678–8681.

(93) Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Voorhis, T. Van; Pentelute, B. L. *Nat. Chem.* **2016**, *8*, 120–128.

(94) Dai, P.; Williams, J. K.; Zhang, C.; Welborn, M.; Shepherd, J. J.; Zhu, T.; Voorhis, T. V.; Hong, M.; Pentelute, B. L. *Sci. Rep.* **2017**, *7* (7954), 1–9.

(95) Evans, E. D.; Pentelute, B. L. *ACS Chem. Biol.* **2018**, *13*, 527–532.

(96) Zhang, C.; Dai, P.; Vinogradov, A. A.; Gates, Z. P.; Pentelute, B. L. *Angew. Chem. Int. Ed.* **2018**, *57* (22), 6459–6463.

(97) Bernard-Gauthier, V.; Wängler, C.; Schirrmacher, E.; Kostikov, A.; Jurkschat, K.; Wängler, B.; Schirrmacher, R. *Biomed Res. Int.* **2014**, *2014*, 1–20.

(98) Verhoog, S.; Kee, C. W.; Wang, Y.; Khotavivattana, T.; Wilson, T. C.; Kersemans, V.; Smart, S.; Tredwell, M.; Davis, B. G.; Gouverneur, V. *J. Am. Chem. Soc.* **2018**, *140*, 1572–1575.

(99) Boutureira, O.; Martínez-Sáez, N.; Brindle, K. M.; Neves, A. A.; Corzana, F.; Bernardes, G. J. L. *Chem. Eur. J.* **2017**, *23* (27), 6483–6489.

(100) Lee, B.; Sun, S.; Jiménez-Moreno, E.; Neves, A. A.; Bernardes, G. J. L. *Bioorg. Med. Chem.* **2018**, *26* (11), 3060–3064.

(101) Becker, J. M.; Wilchek, M. *BBA-Gen. Subjects* **1972**, *264* (1), 165–170.

(102) Staros, J. V.; Wright, R. W.; Swingle, D. M. *Anal. Biochem.* **1986**, *156* (1), 220–222.

(103) Traut, R. R.; Bollen, A.; Sun, T.-T.; Hershey, J. W. B.; Sundberg, J.; Pierce, L. R. *Biochemistry* **1973**, *12* (17), 3266–3273.

(104) Wofsy, L.; Singer, S. J. *Biochemistry* **1963**, *2* (1), 104–116.

(105) Lee, Y. C.; Stowell, C. P.; Krantz, M. J. *Biochemistry* **1976**, *15* (18), 3956–3963.

(106) Annunziato, M. E.; Patel, U. S.; Ranade, M.; Palumbo, P. S. *Bioconjugate Chem.* **1993**, *4* (3), 212–218.

(107) Riggs, J. L.; Seiwald, R. J.; Burckhalter, J. H.; Downs, C. M.; Metcalf, T. G. *Am. J.*

Pathol. **1958**, *34* (6), 1081–1097.

(108) Tuls, J.; Geren, L.; Millett, F. *J. Biol. Chem.* **1989**, *264* (28), 16421–16425.

(109) Rosa-Neto, P.; Wängler, B.; Iovkova, L.; Boening, G.; Reader, A.; Jurkschat, K.; Schirrmacher, E. *ChemBioChem* **2009**, *10* (8), 1321–1324.

(110) Hartley, B. S.; Massey, V. *Biochim. Biophys. Acta* **1956**, *21* (1), 58–70.

(111) Lefevre, C.; Kang, H. C.; Haugland, R. P.; Malekzadeh, N.; Arttamangkul, S.; Haugland, R. P. *Bioconjugate Chem.* **1996**, *7* (4), 482–489.

(112) Evangelista, R. A.; Pollak, A.; Allore, B.; Templeton, E. F.; Morton, R. C.; Diamandis, E. P. *Clin. Biochem.* **1988**, *21* (3), 173–178.

(113) Herzig, D. J.; Rees, A. W.; Day, R. A. *Biopolymers* **1964**, *2* (4), 349–360.

(114) Wurm, F.; Dingels, C.; Frey, H.; Klok, H.-A. *Biomacromolecules* **2012**, *13* (4), 1161–1171.

(115) Wurm, F.; Steinbach, T.; Klok, H.-A. *Chem. Commun.* **2013**, *49* (71), 7815.

(116) Gavrilyuk, J. I.; Wuellner, U.; Barbas, C. F. *Bioorg. Med. Chem. Lett.* **2009**, *19* (5), 1421–1424.

(117) Gavrilyuk, J. I.; Wuellner, U.; Salahuddin, S.; Goswami, R. K.; Sinha, S. C.; Barbas, C. F. *Bioorg. Med. Chem. Lett.* **2009**, *19* (14), 3716–3720.

(118) Nanna, A. R.; Li, X.; Walseng, E.; Pedzisa, L.; Goydel, R. S.; Hymel, D.; Burke, T. R.; Roush, W. R.; Rader, C. *Nat. Commun.* **2017**, *8* (1112), 1–9.

(119) Imai, K.; Watanabe, Y. *Anal. Chim. Acta* **1981**, *130* (2), 377–383.

(120) Sutton, D. A.; Drewes, S. E.; Welz, U. *Biochem. J.* **1972**, *130* (2), 589–595.

(121) Ladd, D. L.; Snow, R. A. *Anal. Biochem.* **1993**, *210* (2), 258–261.

(122) Marshall, J. J. *Trends Biochem. Sci.* **1978**, *3* (2), 79–83.

(123) Bentley, M. D.; Roberts, M. J.; Harris, J. M. *J. Pharm. Sci.* **1998**, *87* (11), 1446–1449.

(124) Lee, R. T.; Lee, Y. C. *Biochemistry* **1980**, *19* (1), 156–163.

(125) McFarland, J. M.; Francis, M. B. *J. Am. Chem. Soc.* **2005**, *127* (39), 13490–13491.

(126) Degruyter, J. N.; Malins, L. R.; Baran, P. S. *Biochemistry* **2017**, *56*, 3863–3873.

(127) Chaubet, G.; Thoreau, F.; Wagner, A. *Drug Discov. Today Technol.* **2018**, *30*, 21–26.

(128) Akkapeddi, P.; Azizi, S.-A.; Freedy, A. M.; Cal, P. M. S. D.; Gois, P. M. P.; Bernardes, G. J. L. *Chem. Sci.* **2016**, *7* (5), 2954–2963.

(129) Lee, H. G.; Lautrette, G.; Pentelute, B. L.; Buchwald, S. L. *Angew. Chem. Int. Ed.* **2017**, *56* (12), 3177–3181.

(130) Matos, M. J.; Oliveira, B. L.; Martínez-Sáez, N.; Guerreiro, A.; Cal, P. M. S. D.; Bertoldo, J.; Maneiro, M.; Perkins, E.; Howard, J.; Deery, M. J.; Chalker, J. M.; Corzana, F.; Jiménez-Osés, G.; Bernardes, G. J. L. *J. Am. Chem. Soc.* **2018**, *140* (11), 4004–4017.

(131) Isom, D. G.; Castañeda, C. A.; Cannon, B. R.; García-Moreno, B. *P. Natl. A. Sci. USA.* **2011**, *108* (13), 5260–5265.

(132) Hacker, S. M.; Backus, K. M.; Lazear, M. R.; Forli, S.; Correia, B. E.; Cravatt, B. F. *Nat. Chem.* **2017**, *9* (12), 1181–1190.

(133) Choi, S.; Connelly, S.; Reixach, N.; Wilson, I. A.; Kelly, J. W. *Nat. Chem. Biol.* **2010**, *6* (2), 133–139.

(134) Asano, S.; Patterson, J. T.; Gaj, T.; Barbas, C. F. *Angew. Chem. Int. Ed.* **2014**, *53* (44), 11783–11786.

(135) Foettinger, A.; Melmer, M.; Leitner, A.; Lindner, W. *Bioconjugate Chem.* **2007**, *18* (5), 1678–1683.

(136) Antos, J. M.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 10256–10257.

(137) Popp, B. V; Ball, Z. T. *J. Am. Chem. Soc.* **2010**, *132* (19), 6660–6662.

(138) Schischko, A.; Ren, H.; Kaplaneris, N.; Ackermann, L. *Angew. Chem. Int. Ed.* **2017**, *56* (6), 1576–1580.

(139) Seki, Y.; Ishiyama, T.; Sasaki, D.; Abe, J.; Sohma, Y.; Oisaki, K.; Kanai, M. *J. Am. Chem. Soc.* **2016**, *138* (34), 10798–10801.

(140) Uchida, K.; Stadtman, E. R. *P. Natl. A. Sci. USA.* **1992**, *89* (10), 4544–4548.

(141) Li, X.; Ma, H.; Dong, S.; Duan, X.; Liang, S. *Talanta* **2004**, *62* (2), 367–371.

(142) Jia, S.; Chang, C. Bioinspired Thiophosphorodichloridate Reagents for Chemoselective Histidine Bioconjugation https://chemrxiv.org/articles/Bioinspired_Thiophosphorodichloridate_Reagents_for_C

hemoselective_Histidine_Bioconjugation/7250201 (accessed Jan 8, 2019).

(143) Neel S. Joshi; Leanna R. Whitaker; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126* (49), 15942–15943.

(144) Ban, H.; Gavrilyuk, J.; Barbas, C. F. *J. Am. Chem. Soc.* **2010**, *132* (5), 1523–1525.

(145) Alvarez-Dorta, D.; Thobie-Gautier, C.; Croyal, M.; Bouzelha, M.; Mével, M.; Deniaud, D.; Boujtita, M.; Gouin, S. G. *J. Am. Chem. Soc.* **2018**, *140* (49), 17120–17126.

(146) Allan, C.; Kosar, M.; Burr, C. V.; Mackay, C. L.; Duncan, R. R.; Hulme, A. N. *ChemBioChem* **2018**, *19* (23), 2443–2447.

(147) Ohata, J.; Miller, M. K.; Mountain, C. M.; Vohidov, F.; Ball, Z. T. *Angew. Chem. Int. Ed.* **2018**, *57* (11), 2827–2830.

(148) Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem. Int. Ed.* **2006**, *45* (32), 5307–5311.

(149) Scheck, R. A.; Dedeo, M. T.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc.* **2008**, *130* (35), 11762–11770.

(150) Obermeyer, A. C.; Jarman, J. B.; Francis, M. B. *J. Am. Chem. Soc.* **2014**, *136*, 9572–9579.

(151) Chen, D.; Disotuar, M. M.; Xiong, X.; Wang, Y.; Chou, D. H.-C. *Chem. Sci.* **2017**, *8*, 2717–2722.

(152) Bloom, S.; Liu, C.; Kölmel, D. K.; Qiao, J. X.; Zhang, Y.; Poss, M. A.; Ewing, W. R.; MacMillan, D. W. C. *Nat. Chem.* **2017**, *10* (2), 205–211.

(153) Valley, C. C.; Cembran, A.; Perlmutter, J. D.; Lewis, A. K.; Labello, N. P.; Gao, J.; Sachs, J. N. *J. Biol. Chem.* **2012**, *287* (42), 34979–34991.

(154) Cirino, P. C.; Tang, Y.; Takahashi, K.; Tirrell, D. A.; Arnold, F. H. *Biotechnol. Bioeng.* **2003**, *83* (6), 729–734.

(155) Ferla, M. P.; Patrick, W. M. *Microbiology* **2014**, *160*, 1571–1584.

(156) Struck, A.-W.; Thompson, M. L.; Wong, L. S.; Micklefield, J. *ChemBioChem* **2012**, *13* (18), 2642–2655.

(157) Kim, G.; Weiss, S. J.; Levine, R. L. *BBA-Gen. Subjects* **2014**, *1840* (2), 901–905.

(158) Martínez, Y.; Li, X.; Liu, G.; Bin, P.; Yan, W.; Más, D.; Valdivié, M.; Hu, C.-A. A.;

Ren, W.; Yin, Y. *Amino Acids* **2017**, *49*, 2091–2098.

(159) Fu, X.; Adams, Z.; Liu, R.; Hepowit, N. L.; Wu, Y.; Bowmann, C. F.; Moskovitz, J.; Maupin-Furlow, J. A. *MBio* **2017**, *8* (5), e01169-17.

(160) Bigelow D J; Squier, T. C. *Biochim. Biophys. Acta* **2005**, *1703*, 121–134.

(161) Hoffman, R. M. *BBA-Rev. Cancer* **1984**, *738*, 49–87.

(162) Birben, E.; Sahiner, U. M.; Sackesen, C.; Erzurum, S.; Kalayci, O. *World Allergy Organ. J.* **2012**, *5* (1), 9–19.

(163) van der Westhuyzen, J. *Nutr. Cancer* **1985**, *7* (3), 179–183.

(164) Patra, R. .; Swarup, D.; Dwivedi, S. . *Toxicology* **2001**, *162* (2), 81–88.

(165) Jori, G.; Galiazzo, G.; Marzotto, A.; Scoffone, E. *BBA-Protein Struct. M.* **1968**, *154* (1), 1–9.

(166) Jori, G.; Galiazzo, G.; Marzotto, A.; Scoffone, E. *J. Biol. Chem.* **1968**, *243*, 4272–4278.

(167) Misani, F.; Fair, T. W.; Reiner, L. *J. Am. Chem. Soc.* **1951**, *73* (1), 459–461.

(168) Whitehead, J. K.; Bentley, H. R. *J. Chem. Soc.* **1952**, *0* (0), 1572.

(169) Brill, A. S.; Weinryb, I. *Biochemistry* **1967**, *6* (11), 3528–3535.

(170) Toennies, G. *J. Biol. Chem.* **1940**, *132*, 455–455.

(171) Link, T. P.; Stark, G. R. *J. Biol. Chem.* **1968**, *243*, 1082–1088.

(172) Toennies, G.; Kolb, J. J. *J. Am. Chem. Soc.* **1945**, *67* (5), 849–851.

(173) Gerd Gundlach, H.; Moore, S.; Stein, W. H. *J. Biol. Chem.* **1959**, *234* (7), 1761–1764.

(174) Mondal, J.; Tiwary, P.; Berne, B. J. *J. Am. Chem. Soc.* **2016**, *138* (13), 4608–4615.

(175) Gundlach, H. G.; Steip, W. H.; Moore, S. *J. Biol. Chem.* **1959**, *234* (7), 1754–1760.

(176) Bradshaw, R. A.; Robinson, G. W.; Hass, G. M.; Hill, R. L. *J. Biol. Chem.* **1969**, *244*, 1755–1763.

(177) Lawson, W. B.; Schramm, H. J. *J. Am. Chem. Soc.* **1962**, *84* (10), 2017–2018.

(178) Lawson, W. B.; Schramm, H.-J. *Biochemistry* **1965**, *4* (3), 377–385.

(179) Jones, J. B.; Hysert, D. W. *Can. J. Chem.* **1971**, *49*, 3012–3019.

(180) Naider, F.; Bohak, Z. *Biochemistry* **1972**, *11* (17), 3208–3211.

(181) Naider, F.; Bohak, Z.; Yariv, J. *Biochemistry* **1972**, *11* (17), 3202–3207.

(182) Doi, J. T.; Luehr, G. W. *Tetrahedron Lett.* **1985**, *26* (50), 6143–6146.

(183) Vithayathil, P. J.; Richards, F. M. *J. Biol. Chem.* **1960**, *235* (8), 2343–2351.

(184) Gorent, H. J.; Barnard, E. A. *Biochemistry* **1970**, *9* (4), 959–973.

(185) Landis, B. H.; Berliner, L. J. *J. Am. Chem. Soc.* **1980**, *102* (16), 5350–5354.

(186) Kramer, J. R.; Deming, T. J. *Biomacromolecules* **2012**, *13*, 1719–1723.

(187) Kramer, J. R.; Deming, T. J. *Chem. Commun.* **2013**, *49* (49), 5144–5146.

(188) Gharakhanian, E. G.; Deming, T. J. *Chem. Commun* **2016**, *52*, 5336–5339.

(189) Gross, E.; Witkop, B. *J. Am. Chem. Soc.* **1961**, *83* (6), 1510–1511.

(190) Gross, E.; Witkop, B. *J. Biol. Chem.* **1962**, *237* (6), 1856–1860.

(191) Andreev, Y. A.; Kozlov, S. A.; Vassilevski, A. A.; Grishin, E. V. *Anal. Biochem.* **2010**, *407* (1), 144–146.

(192) Lin, S.; Yang, X.; Jia, S.; Weeks, A. M.; Hornsby, M.; Lee, P. S.; Nichiporuk, R. V; Iavarone, A. T.; Wells, J. A.; Toste, F. D.; Chang, C. J. *Science* **2017**, *355*, 597–602.

(193) Kramer, J. R.; Schmidt, N. W.; Mayle, K. M.; Kamei, D. T.; Wong, G. C. L.; Deming, T. J. *ACS Cent. Sci.* **2015**, *1*, 83–88.

(194) Maaskant, R. V.; Roelfes, G. *ChemBioChem* **2019**, *20* (1), 57–61.

(195) Zhdankin, V. V; Stang, P. J. *Chem. Rev.* **2008**, *108*, 5299–5368.

(196) Merritt, E.; Olofsson, B. *Angew. Chem. Int. Ed.* **2009**, *48* (48), 9052–9070.

(197) Perkins, C. W.; Martin, J. C.; Arduengo, A. J.; Lau, W.; Alegria, A.; Kochi, J. K. *J. Am. Chem. Soc* **1980**, *102* (26), 7753–7759.

(198) Malmgren, J.; Santoro, S.; Jalalian, N.; Himo, F.; Olofsson, B. *Chem. Eur. J.* **2013**, *19* (31), 10334–10342.

(199) Stang, P. J.; Zhdankin, V. V. *Chem. Rev.* **1996**, *96*, 1123–1178.

(200) Yoshimura, A.; Zhdankin, V. V. *Chem. Rev.* **2016**, *116*, 3328–3435.

(201) Zhdankin, V. V.; Stang, P. J. *Chem. Rev.* **2002**, *102* (7), 2523–2584.

(202) Okuyama, T.; Takino, T.; Sueda, T.; Ochiai, M. *J. Am. Chem. Soc.* **1995**, *117* (12), 3360–3367.

(203) R Crowder, B. J.; Glover, E. E.; Grundon, M. F.; Kaempfen, H. X. *J. Chem. Soc.* **1963**, 4578–4585.

(204) Chan, L.; Mcnally, A.; Toh, Q. Y.; Mendoza, A.; Gaunt, M. J. *Chem. Sci.* **2015**, *6*, 1277–1281.

(205) Carroll, M. A.; Nairne, J.; Woodcraft, J. L. *J. Labelled Compd. Rad.* **2007**, *50* (5–6), 452–454.

(206) Martín-Santamaría, S.; Carroll, M. A.; Carroll, C. M.; Carter, C. D.; Pike, V. W.; Rzepa, H. S.; Widdowson, D. A. *Chem. Commun.* **2000**, 649–650.

(207) Yan, J.; Hu, W.; Rao, G. *Synthesis* **2006**, *6*, 943–945.

(208) Yan, J.; Hu, W.; Zhou, W. *Synth. Commun.* **2006**, *36*, 2097–2102.

(209) Yan, J.; Zhu, M.; Zhou, Z. *Eur. J. Org. Chem.* **2006**, *2006* (9), 2060–2062.

(210) Carroll, M. A.; Wood, R. A. *Tetrahedron* **2007**, *63*, 11349–11354.

(211) Krief, A.; Dumont, W.; Robert, M. *Chem. Commun.* **2005**, 2167–2168.

(212) Krief, A.; Dumont, W.; Robert, M.; Professor Atta-ur-Rahman, T. *Synlett* **2006**, *16*, 2601–2604.

(213) Krief, A.; Dumont, W.; Robert, M. *Synlett* **2006**, *3*, 484–486.

(214) Matoušek, V.; Václavík, J.; Hájek, P.; Charpentier, J.; Blastik, Z. E.; Pietrasiak, E.; Budinská, A.; Togni, A.; Beier, P. *Chem. Eur. J.* **2016**, *22* (1), 417–424.

(215) Václavík, J.; Zschoche, R.; Klimánková, I.; Matoušek, V.; Beier, P.; Hilvert, D.; Togni, A. *Chem. Eur. J.* **2017**, *23* (27), 6490–6494.

(216) Frei, R.; Jérô Me Waser, J. *J. Am. Chem. Soc* **2013**, *135*, 9620–9623.

(217) Abegg, D.; Frei, R.; Cerato, L.; Prasad Hari, D.; Wang, C.; Waser, J.; Adibekian, A. *Angew. Chem. Int. Ed.* **2015**, *54* (37), 10852–10857.

(218) Tolnai, G. L.; Brand, J. P.; Waser, J. *Beilstein J. Org. Chem* **2016**, *12*, 745–749.

(219) Hansen, M. B.; Hubálek, F.; Skrydstrup, T.; Hoeg-Jensen, T. *Chem. Eur. J.* **2016**, *22* (5), 1572–1576.

(220) Weiss, R.; Seubert, J.; Hampel, F. *Angew. Chem. Int. Ed.* **1994**, *33* (19), 1952–1953.

(221) Sahu, S.; Sahoo, R.; Patel, S.; Mishra, B. K. *J. Sulfur Chem.* **2011**, *32* (2), 171–197.

(222) Vogler, T.; Studer, A. *Synthesis* **2008**, *2008* (13), 1979–1993.

(223) Cvetkovic, R. S.; Plosker, G. L. *Drugs* **2007**, *67* (6), 935–954.

(224) Leuchte, H. H.; Baezner, C.; Baumgartner, R. A.; Bevec, D.; Bacher, G.; Neurohr, C.; Behr, J. *Eur. Respir. J.* **2008**, *32*, 1289–1294.

(225) Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatisky, M. *Drug Discov. Today* **2010**, *15*, 40–56.

(226) Saag, K. G.; Shane, E.; Boonen, S.; Marín, F.; Donley, D. W.; Taylor, K. A.; Dalsky, G. P.; Marcus, R. *N. Engl. J. Med.* **2007**, *357*, 2028–2067.

(227) Cottrell, J. S. *J. Proteomics* **2011**, *74*, 1842–1851.

(228) Matthiesen, R.; Bunkenborg, J. In *Mass Spectrometry Data Analysis in Proteomics*; Matthiesen, R., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2013; Vol. 1007, pp 1–45.

(229) Greenfield, N. J. *Nat. Protoc.* **2006**, *1* (6), 2876–2890.

(230) Baldwin, A. D.; Kiick, K. L. *Bioconjugate Chem.* **2011**, *22* (10), 1946–1953.

(231) Lewis, M. R.; Shively, J. E. *Bioconjugate Chem.* **1998**, *9* (1), 72–86.

(232) Lapko, V. N.; Smith, D. L.; Smith, J. B. *J. Mass Spectrom.* **2000**, *35* (4), 572–575.

(233) Reid, G. E.; Roberts, K. D.; Simpson, R. J.; O'Hair, R. A. J. *J. Am. Soc. Mass Spectr.* **2005**, *16* (7), 1131–1150.

(234) Schneider, C.; LaFortune, J. H. W.; Melen, R. L.; Stephan, D. W. *Dalt. Trans.* **2018**, *47* (36), 12742–12749.

(235) Doyle, M. P.; Duffy, R.; Ratnikov, M.; Zhou, L. *Chem. Rev.* **2010**, *110*, 704–724.

(236) Ciszewski, Ł. W.; Rybicka-Jasińska, K.; Gryko, D. *Org. Biomol. Chem.* **2019**, *17* (3), 432–448.

(237) Dang, H.-S.; Roberts, B. P. *J. Chem. Soc. Perk. T. 1* **1996**, 769–775.

(238) Horner, L.; Schwarz, H. *Liebigs Ann. Chem.* **1971**, *747* (1), 1–13.

(239) Banga, A. K.; Chien, Y. W. *Int. J. Pharm.* **1988**, *48*, 15–50.

(240) Szychowski, J.; Mahdavi, A.; Hodas, J. J. L.; Bagert, J. D.; Ngo, J. T.; Landgraf, P.; Dieterich, D. C.; Schuman, E. M.; Tirrell, D. A. *J. Am. Chem. Soc.* **2010**, *132* (51), 18351–18360.

(241) Connon, S. J. *Org. Biomol. Chem.* **2007**, *5*, 3407–3417.

(242) Buckler, S. A.; Doll, L.; Lind, F. K.; Epstein, M. *J. Org. Chem.* **1962**, *27* (3), 794–798.

(243) Takagaki, Y.; Gupta, C. M.; Khorana, H. G. *Biochem. Bioph. Res. Co.* **1980**, *95* (2), 589–595.

(244) Prier, C. K.; Rankic, D. A.; Macmillan, D. W. C. *Chem. Rev.* **2013**, *113*, 5322–5363.

(245) Walton, J. C.; Portela-Cubillo, F. In *Encyclopedia of Reagents for Organic Synthesis*; John Wiley & Sons, Ltd: Chichester, UK, 2007.

(246) Hedstrand, D. M.; Kruizinga, W. H.; Kellogg, R. M. *Tetrahedron Lett.* **1978**, *19* (14), 1255–1258.

(247) Zhu, X.-Q.; Liu, Y.-C.; Cheng, J.-P. *J. Org. Chem.* **1999**, *64*, 8980–8981.

(248) Huang, W.; Cheng, X. *Synlett* **2017**, *28*, 148–158.

(249) Li, G.; Chen, R.; Wu, L.; Fu, Q.; Zhang, X.; Tang, Z. *Angew. Chem. Int. Ed.* **2013**, *52* (32), 8432–8436.

(250) Chen, W.; Liu, Z.; Tian, J.; Li, J.; Ma, J.; Cheng, X.; Li, G. *J. Am. Chem. Soc.* **2016**, *138* (38), 12312–12315.

(251) Goldmann, S.; Born, L.; Kazda, S.; Pittel, B.; Schramm, M. *J. Med. Chem.* **1990**, *33* (5), 1413–1418.

(252) Zempleni, J.; Wijeratne, S. S. K.; Hassan, Y. I. *BioFactors* **2009**, *35* (1), 36–46.

(253) Van Bergen, T. J.; Hedstrand, D. M.; Kruizinga, W. H.; Kellogg, R. M. *J. Org. Chem.* **1979**, *44* (26), 4953–4962.

(254) Taylor, M. T.; Nelson, J. E.; Suero, M. G.; Gaunt, M. J. *Nature* **2018**, *562* (7728), 563–568.

(255) Maruani, A.; Richards, D. A.; Chudasama, V. *Org. Biomol. Chem* **2016**, *14*, 6165–6178.

(256) Banerjee, D.; Liu, A. P.; Voss, N. R.; Schmid, S. L.; Finn, M. G. *ChemBioChem* **2010**, *11* (9), 1273–1279.

(257) Stephanopoulos, N.; Tong, G. J.; Hsiao, S. C.; Francis, M. B. *ACS Nano* **2010**, *4* (10), 6014–6020.

(258) Maruani, A.; Smith, M. E. B.; Miranda, E.; Chester, K. A.; Chudasama, V.; Caddick, S. *Nat. Commun.* **2015**, *6* (6645), 1–9.

(259) Joo, C.; Balci, H.; Ishitsuka, Y.; Buranachai, C.; Ha, T. *Annu. Rev. Biochem.* **2008**, *77*, 51–76.

(260) Cecconi, C.; Shank, E. A.; Bustamante, C.; Marqusee, S. *Science* **2005**, *309* (5743), 2057–2060.

(261) Temming, R. P.; Eggermont, L.; van Eldijk, M. B.; van Hest, J. C. M.; van Delft, F. L. *Org. Biomol. Chem.* **2013**, *11* (17), 2772–2779.

(262) Yamaguchi, A.; Matsuda, T.; Ohtake, K.; Yanagisawa, T.; Yokoyama, S.; Fujiwara, Y.; Watanabe, T.; Hohsaka, T.; Sakamoto, K. *Bioconjugate Chem.* **2016**, *27* (1), 198–206.

(263) Ratner, V.; Kahana, E.; Eichler, M.; Haas, E. *Bioconjugate Chem.* **2002**, *13* (5), 1163–1170.

(264) Sonntag, M. H.; Ibach, J.; Nieto, L.; Verveer, P. J.; Brunsveld, L. *Chem. Eur. J.* **2014**, *20* (20), 6019–6026.

(265) Nathani, R. I.; Moody, P.; Chudasama, V.; Smith, M. E. B.; Fitzmaurice, R. J.; Caddick, S. *Chem. Sci.* **2013**, *4*, 3455–3458.

(266) Puthenveetil, S.; Musto, S.; Loganzo, F.; Tumey, L. N.; O'Donnell, C. J.; Graziani, E. *Bioconjugate Chem.* **2016**, *27* (4), 1030–1039.

(267) Crochet, A. P.; Mohiuddin, M. K.; Francis, M. B.; Paavolam C. D. *Biosens. Bioelectron.* **2010**, *26*, 55–61.

(268) Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. *Nature* **2010**, *464* (7287), 441–444.

(269) Wan, W.; Huang, Y.; Wang, Z.; Russell, W. K.; Pai, P.-J.; Russell, D. H.; Liu, W. R. *Angew. Chem. Int. Ed.* **2010**, *49* (18), 3211–3214.

(270) Johnson, D. B. F.; Xu, J.; Shen, Z.; Takimoto, J. K.; Schultz, M. D.; Schmitz, R. J.; Xiang, Z.; Ecker, J. R.; Briggs, S. P.; Wang, L. *Nat. Chem. Biol.* **2011**, *7* (11), 779–786.

(271) Park, H.-S.; Hohn, M. J.; Umehara, T.; Guo, L.-T.; Osborne, E. M.; Benner, J.; Noren, C. J.; Rinehart, J.; Soll, D.; Reppas, N. B.; Emig, C. J.; Bang, D.; Hwang, S. J.; Jewett, M. C.; Jacobson, J. M.; Church, G. M. *Science* **2011**, *333* (6046), 1151–1154.

(272) van Kasteren, S. I.; Kramer, H. B.; Jensen, H. H.; Campbell, S. J.; Kirkpatrick, J.; Oldham, N. J.; Anthony, D. C.; Davis, B. G. *Nature* **2007**, *446* (7139), 1105–1109.

(273) Brustad, E. M.; Lemke, E. A.; Schultz, P. G.; Deniz, A. A. *J. Am. Chem. Soc.* **2008**, *130* (52), 17664–17665.

(274) Simon, M.; Zangemeister-Wittke, U.; Plückthun, A. *Bioconjugate Chem.* **2012**, *23* (2), 279–286.

(275) Moroder, L. *J. Pept. Sci.* **2005**, *11* (4), 187–214.

(276) Kim, J.; Seo, M.-H.; Lee, S.; Cho, K.; Yang, A.; Woo, K.; Kim, H.-S.; Park, H.-S. *Anal. Chem.* **2013**, *85*, 1468–1474.

(277) Strohalm, M.; Hassman, M.; Košata, B.; Kodíček, M. *Rapid Commun. Mass Sp.* **2008**, *22* (6), 905–908.

(278) Ideue, E.; Toma, T.; Shimokawa, J.; Fukuyama, T. *Org. Synth.* **2012**, *89*, 501–509.

(279) Suero, M. G.; Bayle, E. D.; Collins, B. S. L.; Gaunt, M. J. *J. Am. Chem. Soc.* **2013**, *135* (14), 5332–5335.

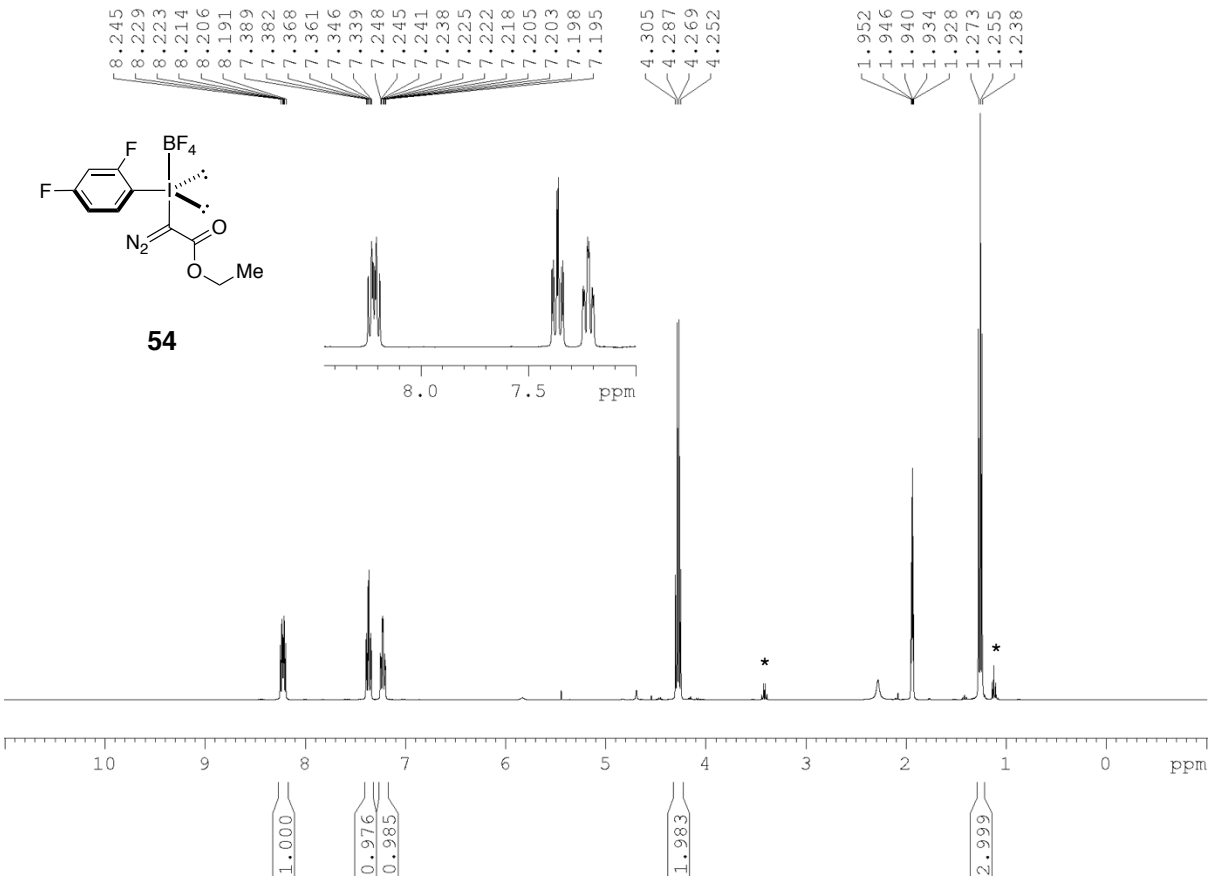
(280) Glunz, P. W.; Mueller, L.; Cheney, D. L.; Ladziata, V.; Zou, Y.; Wurtz, N. R.; Wei, A.; Wong, P. C.; Wexler, R. R.; Priestley, E. S. *J. Med. Chem.* **2016**, *59* (8), 4007–4018.

(281) Lauber, M. B.; Stahl, S. S. *ACS Catal.* **2013**, *3* (11), 2612–2616.

(282) Kim, Y.-W.; Grossmann, T. N.; Verdine, G. L. *Nat. Protoc.* **2011**, *6* (6), 761–771.

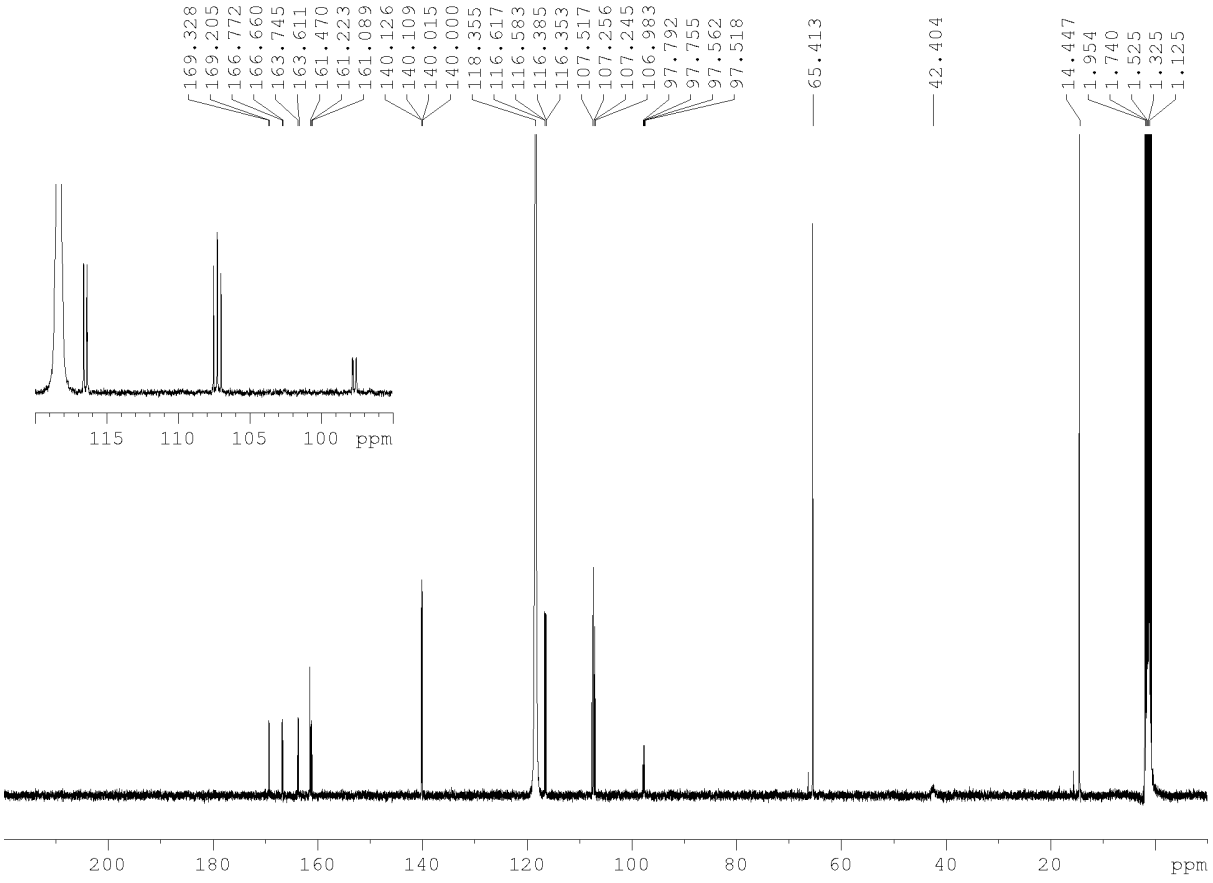
Appendix I

Key spectra

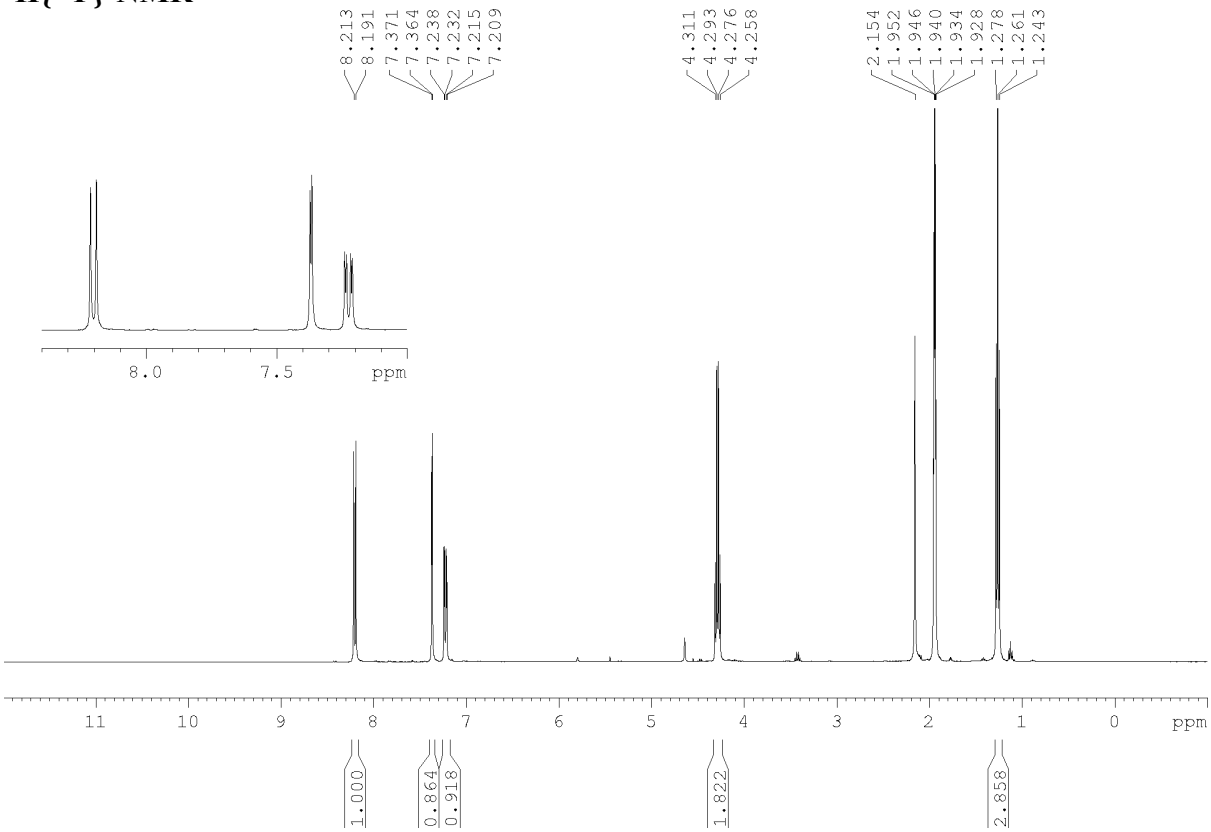


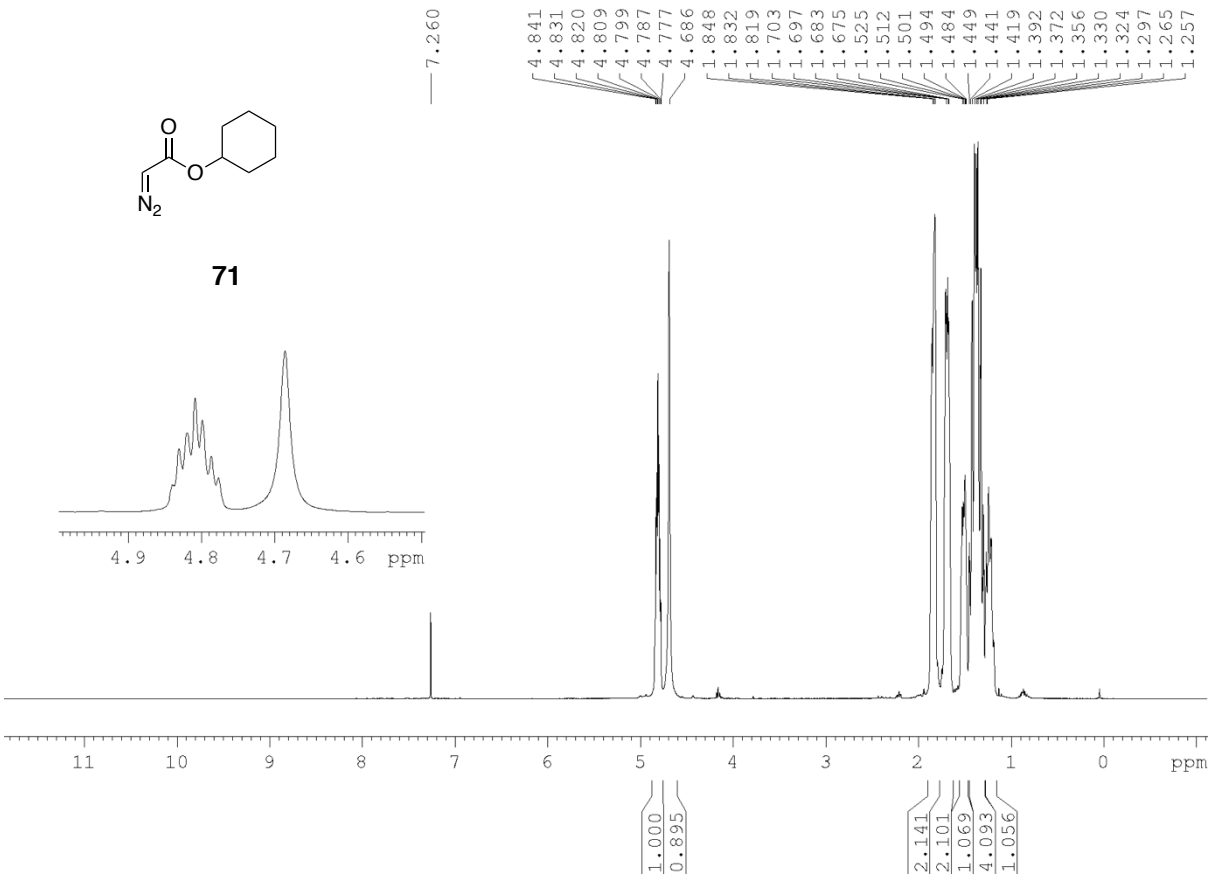
54

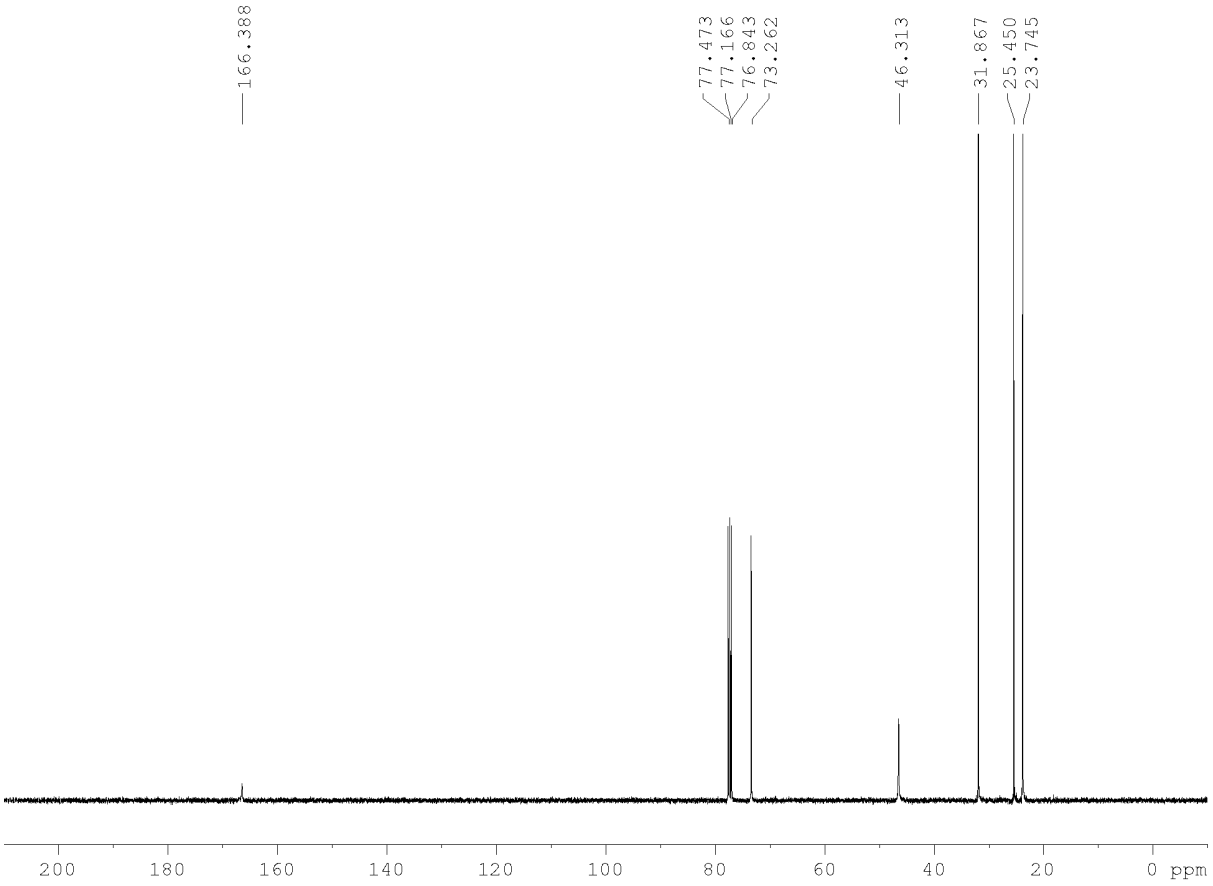
*Impurities

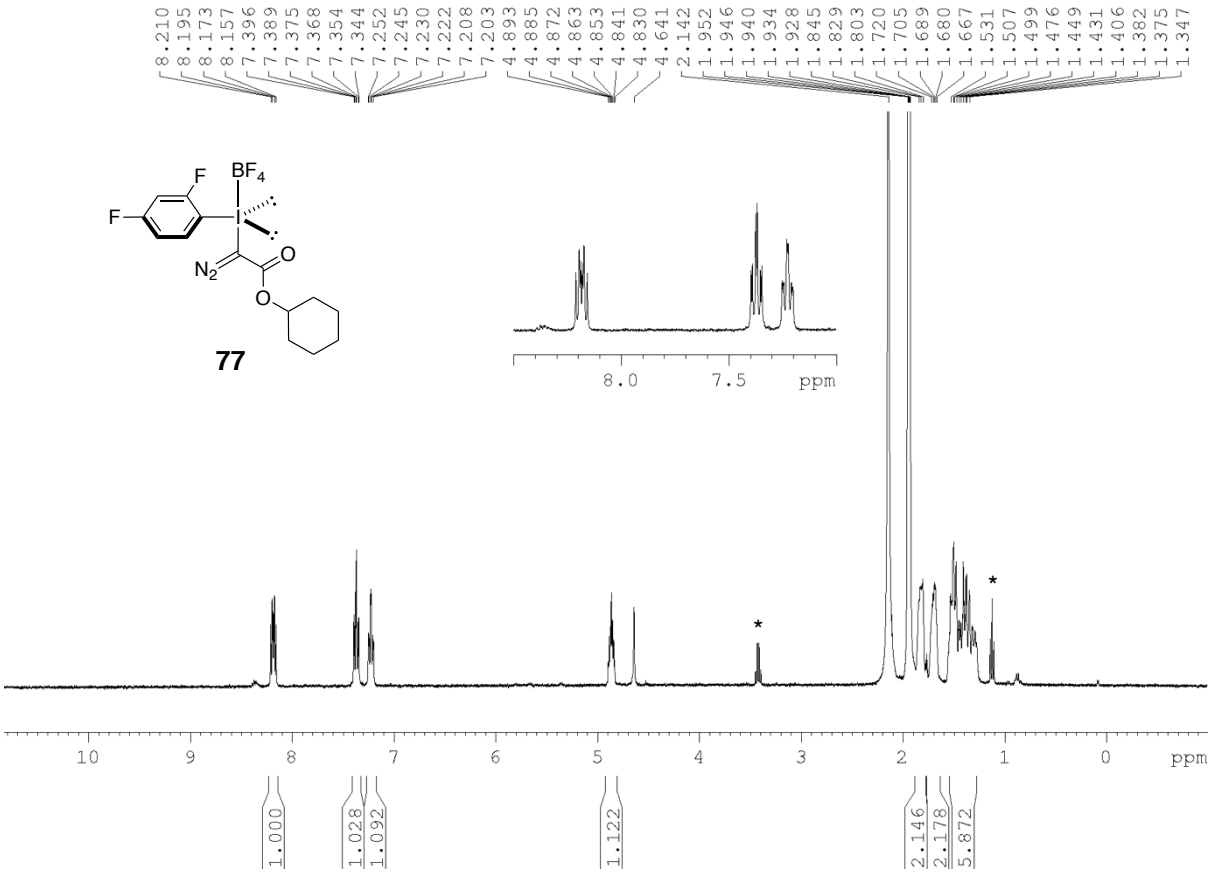


$^{1}H\{^{19}F\}$-NMR

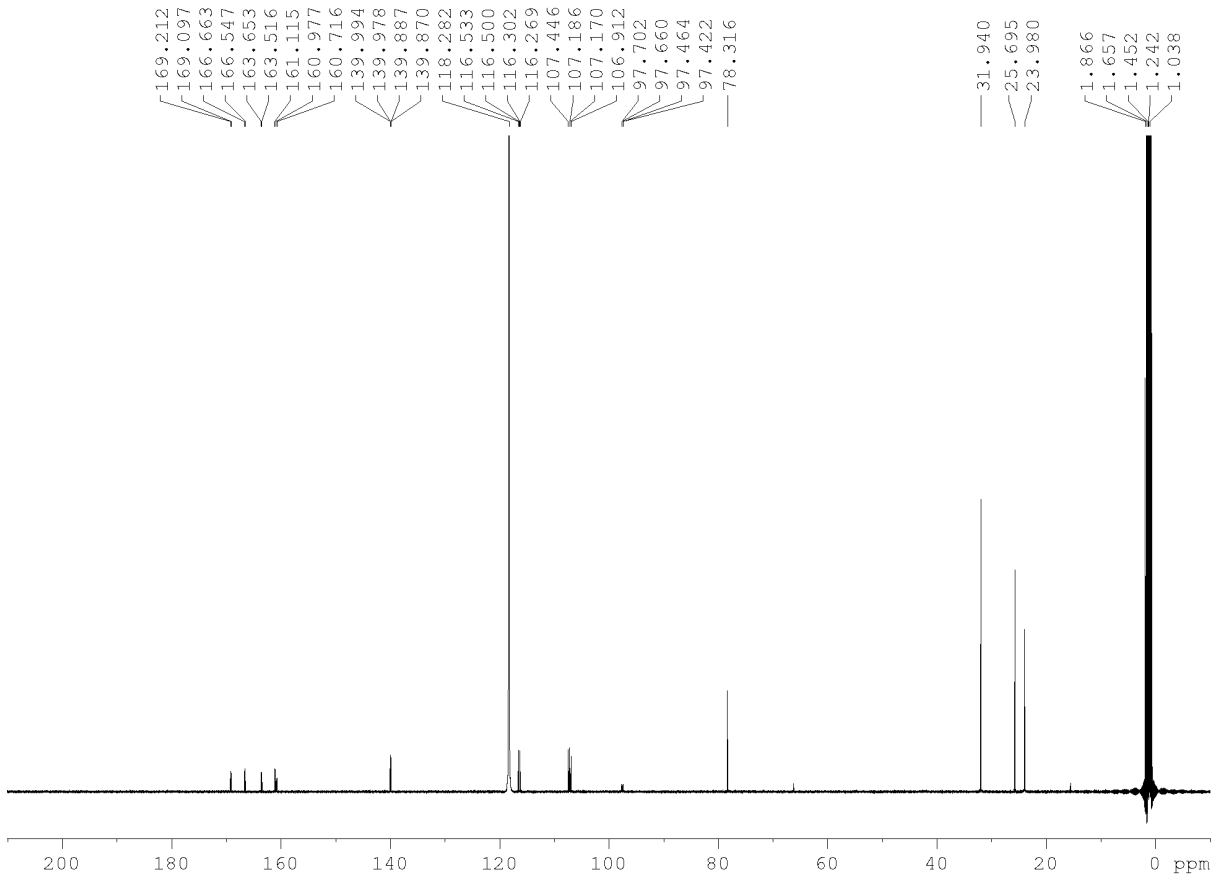

8.213
8.191
7.371
7.364
7.238
7.232
7.215
7.209
4.311
4.293
4.276
4.258
2.154
1.952
1.946
1.940
1.934
1.928
1.278
1.261
1.243
8.0
7.5
ppm
11
10
9
8
7
6
5
4
3
2
1
0
ppm
1.000
0.864
0.918
1.822
2.858

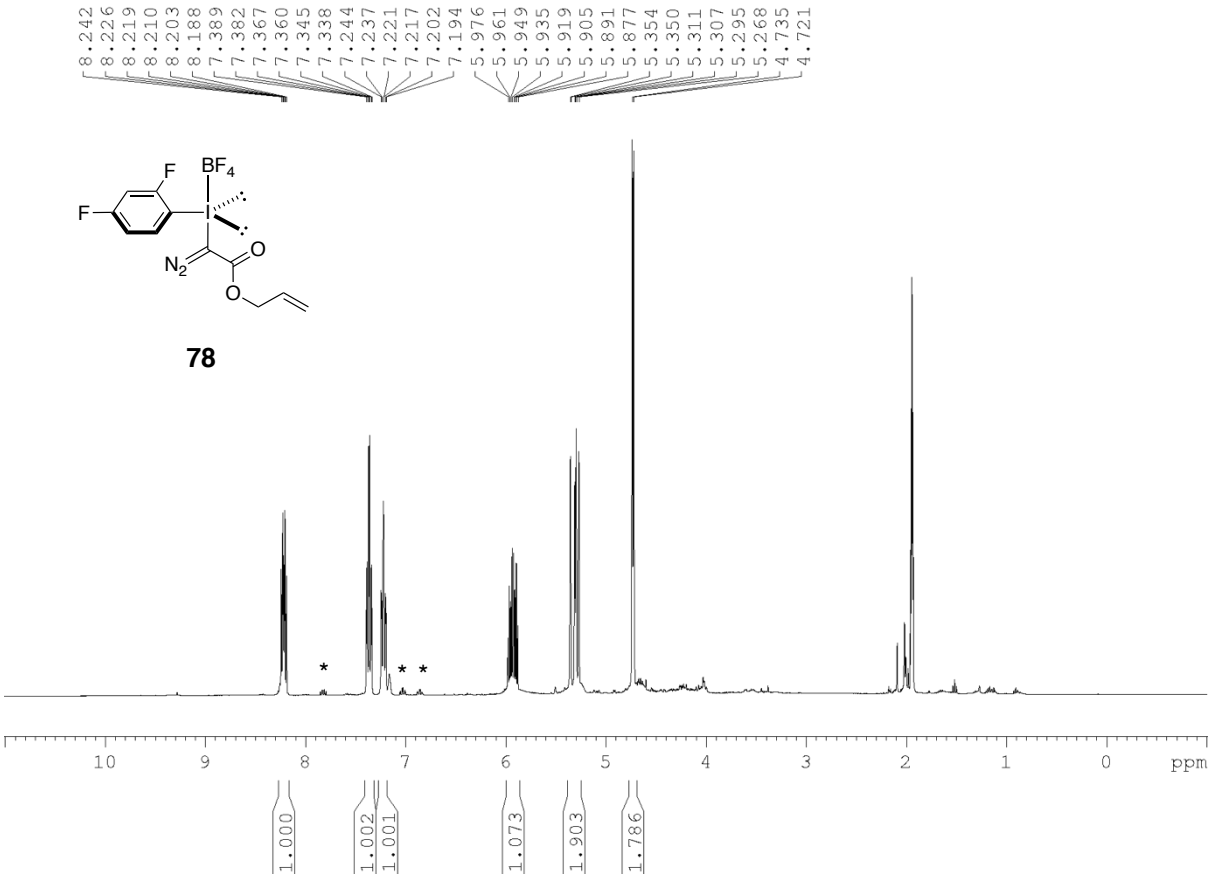

7.260
4.841
4.831
4.820
4.809
4.799
4.787
4.777
4.686
1.848
1.832
1.819
1.703
1.697
1.683
1.675
1.525
1.512
1.501
1.494
1.484
1.449
1.441
1.419
1.392
1.372
1.356
1.330
1.324
1.297
1.265
1.257
O
O
N2
71
4.9
4.8
4.7
4.6
ppm
11
10
9
8
7
6
5
4
3
2
1
0
ppm
1.000
0.895
2.141
2.101
1.069
4.093
1.056


166.388
77.473
77.166
76.843
73.262
46.313
31.867
25.450
23.745
200
180
160
140
120
100
80
60
40
20
0 ppm

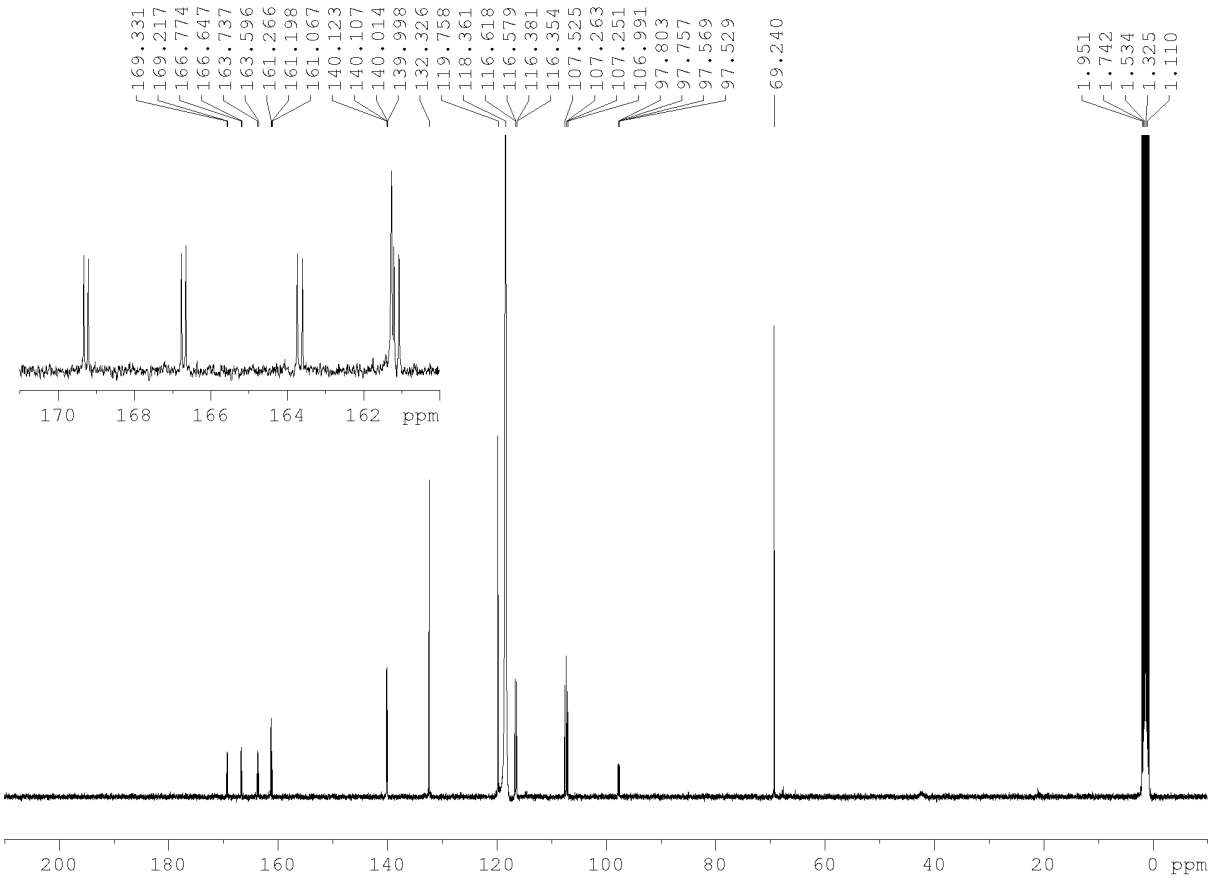

8.210
8.195
8.173
8.157
7.396
7.389
7.375
7.368
7.354
7.344
7.252
7.245
7.230
7.222
7.208
7.203
4.893
4.885
4.872
4.863
4.853
4.841
4.830
4.641
2.142
1.952
1.946
1.940
1.934
1.928
1.845
1.829
1.803
1.720
1.705
1.689
1.680
1.667
1.531
1.507
1.499
1.476
1.449
1.431
1.406
1.382
1.375
1.347
F
BF4
F
I
O
N2
O
77
8.0
7.5
ppm
*
*
10
9
8
7
6
5
4
3
2
1
0
ppm
1.000
1.028
1.092
1.122
2.146
2.178
5.872

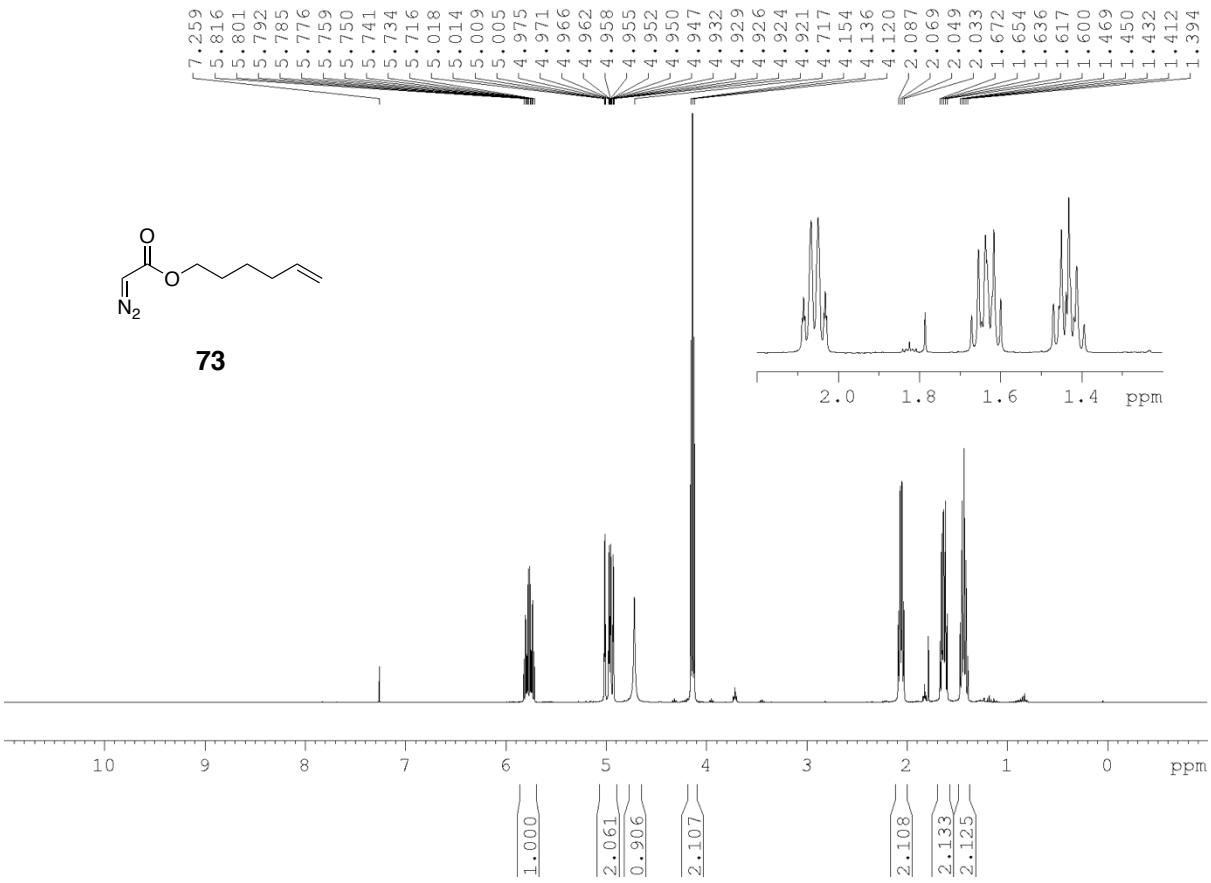
*Impurities

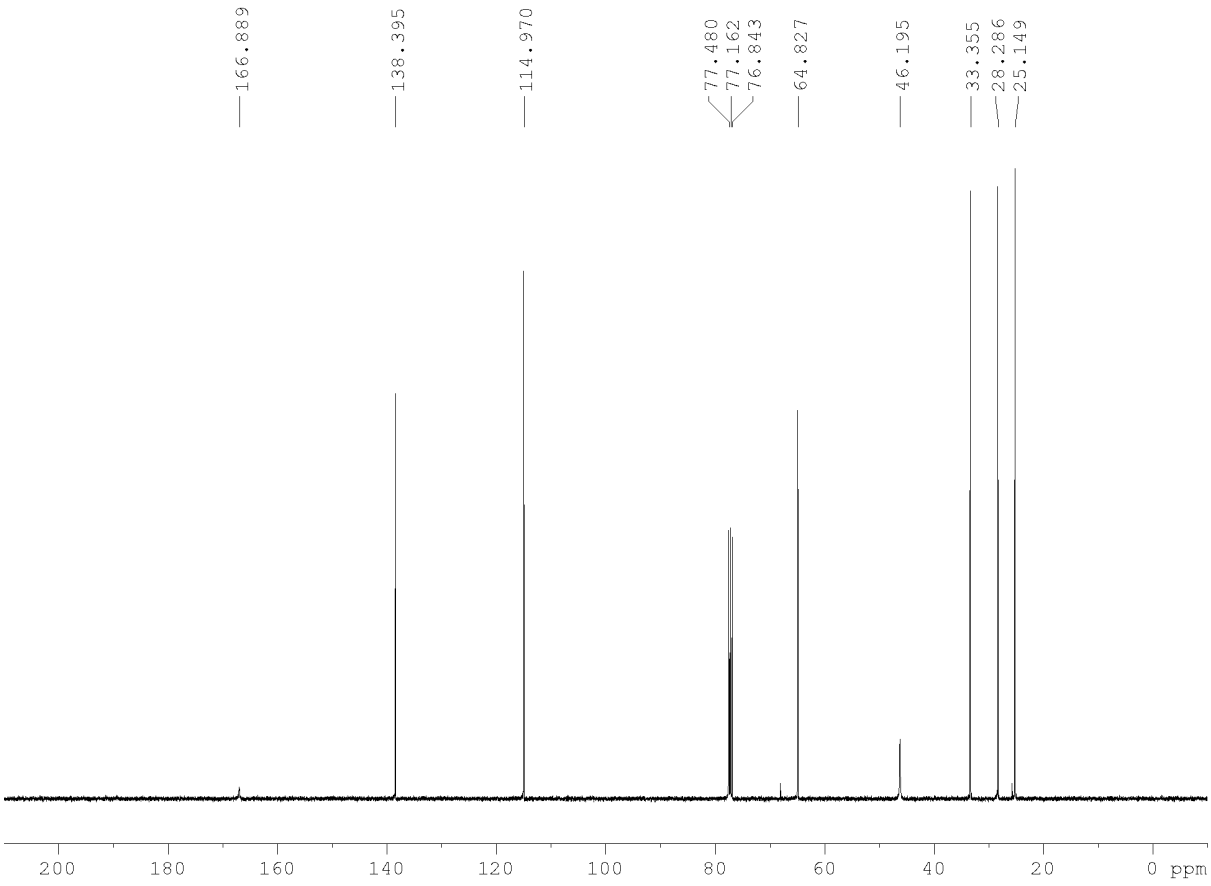

169.212
169.097
166.663
166.547
163.653
163.516
161.115
160.977
160.716
139.994
139.978
139.887
139.870
118.282
116.533
116.500
116.302
116.269
107.446
107.186
107.170
106.912
97.702
97.660
97.464
97.422
78.316
31.940
25.695
23.980
1.866
1.657
1.452
1.242
1.038
200
180
160
140
120
100
80
60
40
20
0
ppm

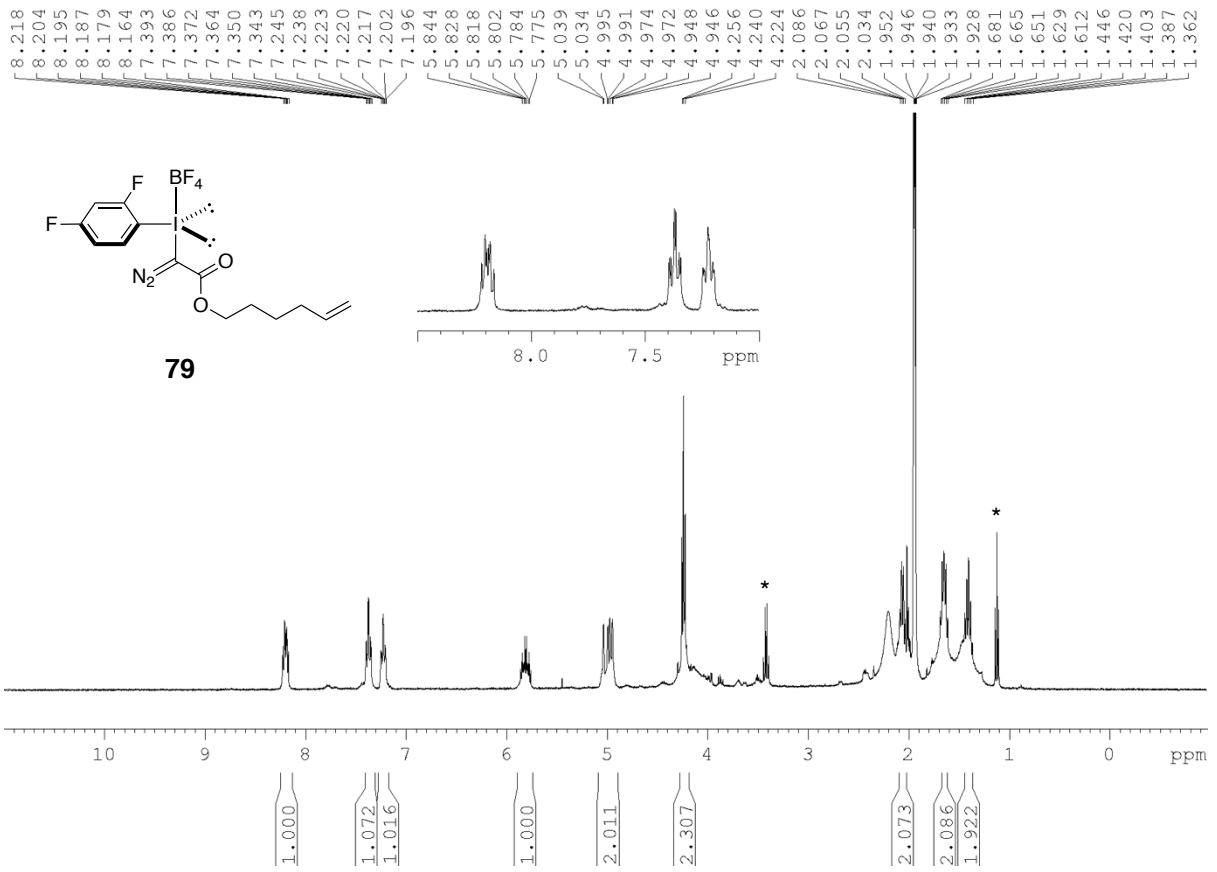

8.242
8.226
8.219
8.210
8.203
8.188
7.389
7.382
7.367
7.360
7.345
7.338
7.244
7.237
7.221
7.217
7.202
7.194
5.976
5.961
5.949
5.935
5.919
5.905
5.891
5.877
5.354
5.350
5.311
5.307
5.295
5.268
4.735
4.721
F
BF4
F
I
N2
O
O
78
*
* *
10
9
8
7
6
5
4
3
2
1
0
ppm
1.000
1.002
1.001
1.073
1.903
1.786

*Impurities


169.331
169.217
166.774
166.647
163.737
163.596
161.266
161.198
161.067
140.123
140.107
140.014
139.998
132.326
119.758
118.361
116.618
116.579
116.381
116.354
107.525
107.263
107.251
106.991
97.803
97.757
97.569
97.529
69.240
1.951
1.742
1.534
1.325
1.110
170
168
166
164
162
ppm
200
180
160
140
120
100
80
60
40
20
0
ppm

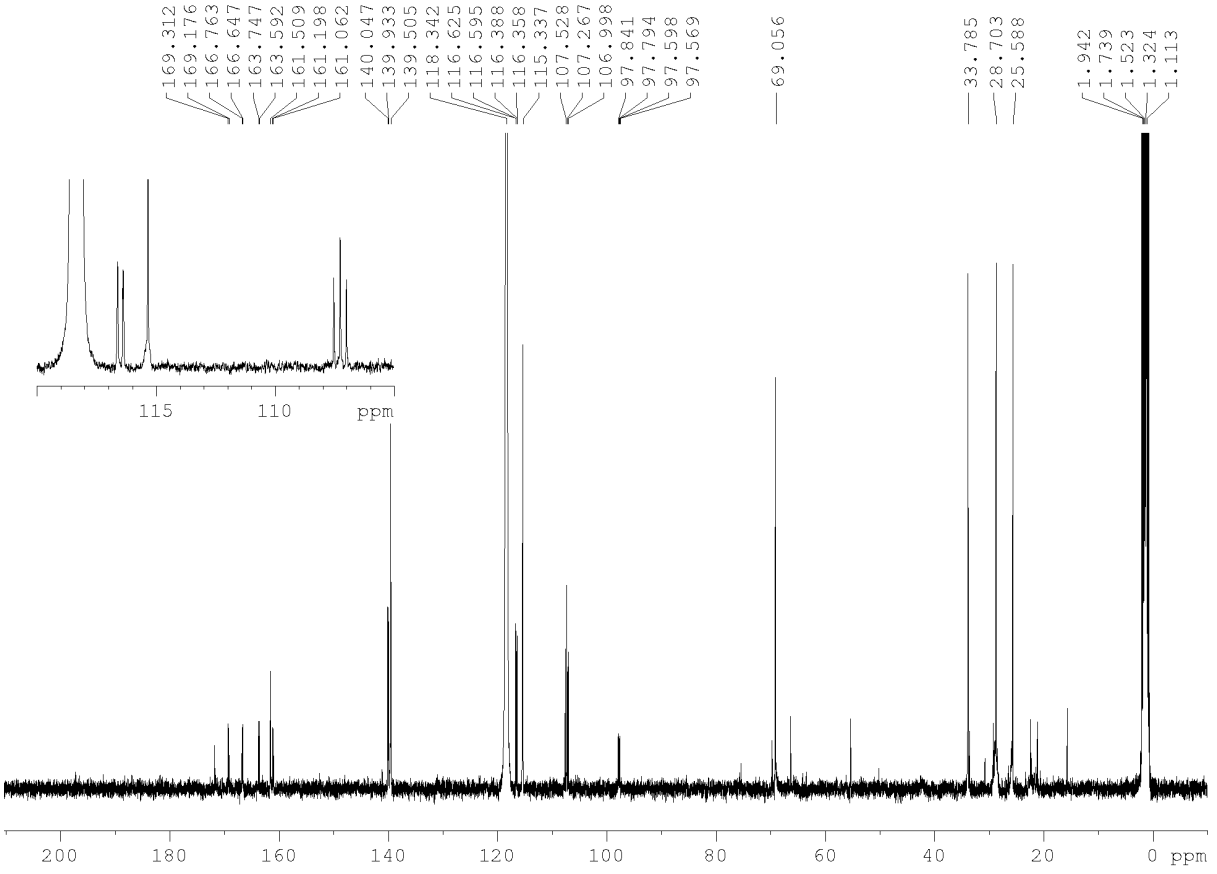

7.259
5.816
5.801
5.792
5.785
5.776
5.759
5.750
5.741
5.734
5.716
5.018
5.014
5.009
5.005
4.975
4.971
4.966
4.962
4.958
4.955
4.952
4.950
4.947
4.932
4.929
4.926
4.924
4.921
4.717
4.154
4.136
4.120
2.087
2.069
2.049
2.033
1.672
1.654
1.636
1.617
1.600
1.469
1.450
1.432
1.412
1.394
O
O
N2
73
2.0
1.8
1.6
1.4
ppm
10
9
8
7
6
5
4
3
2
1
0
ppm
1.000
2.061
0.906
2.107
2.108
2.133
2.125

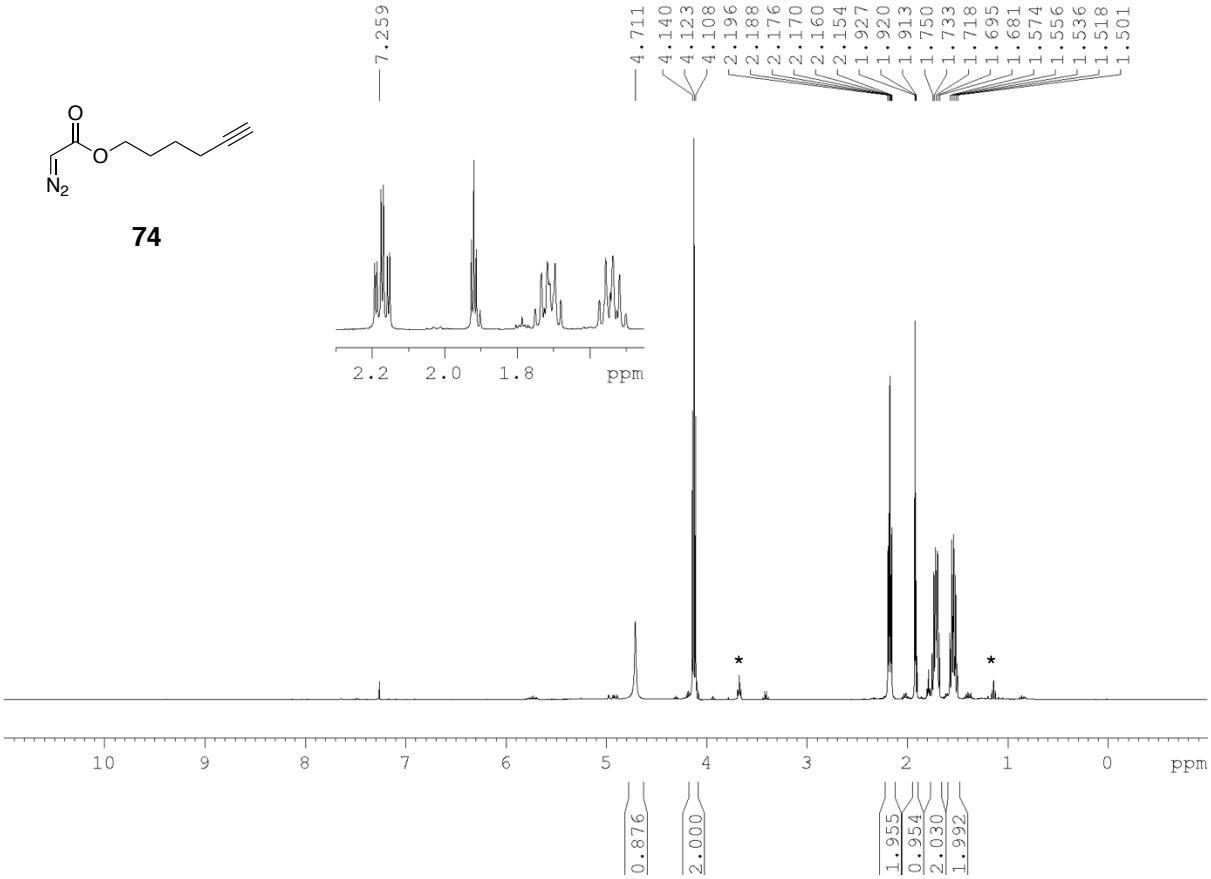

166.889
138.395
114.970
77.480
77.162
76.843
64.827
46.195
33.355
28.286
25.149
200
180
160
140
120
100
80
60
40
20
0 ppm



79

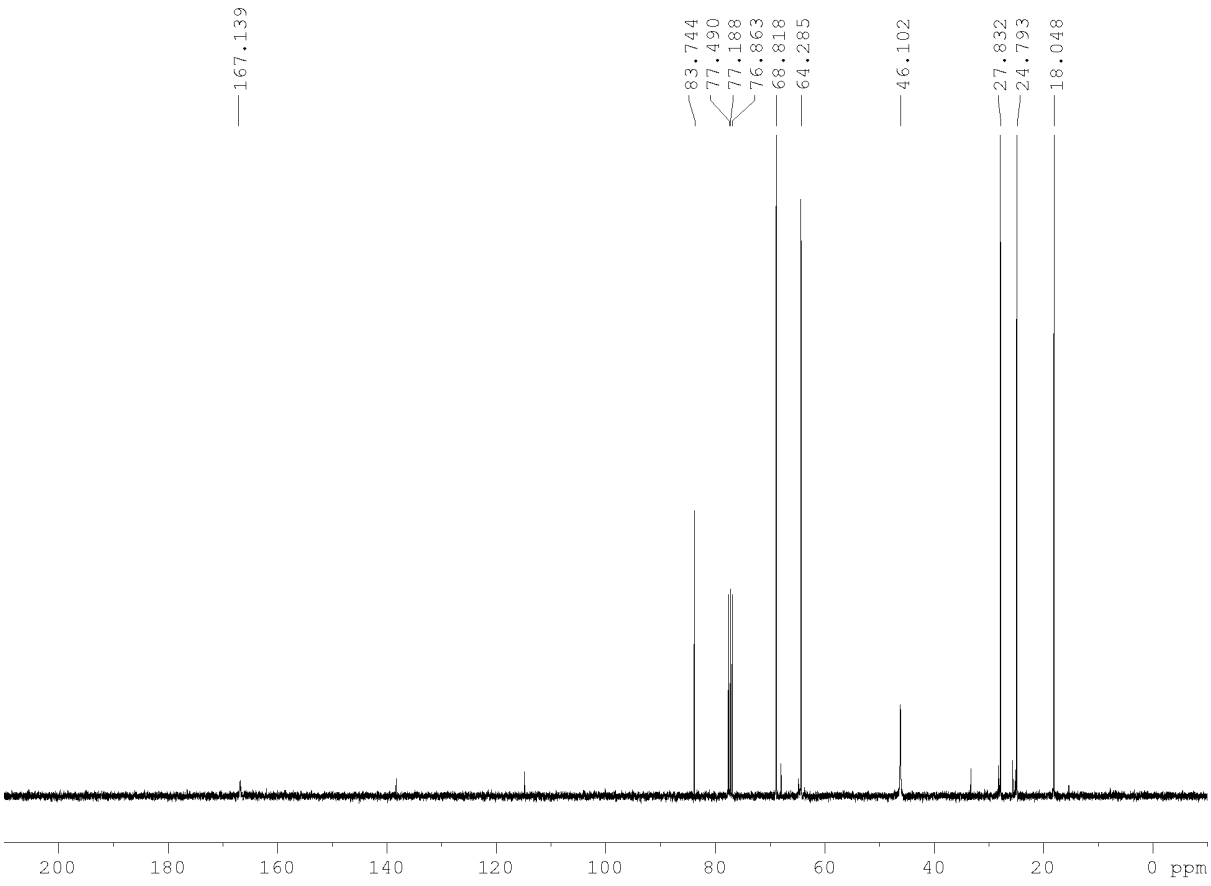
*Impurities

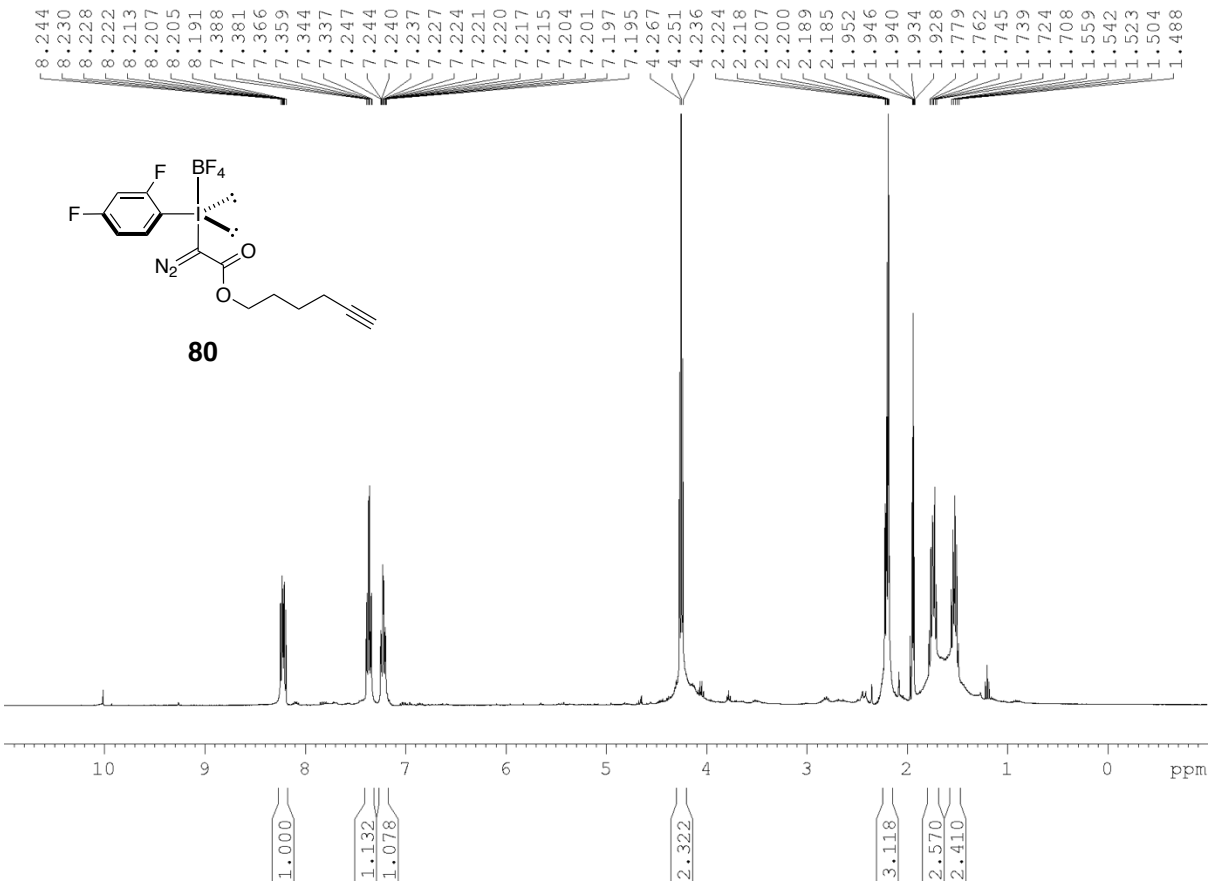




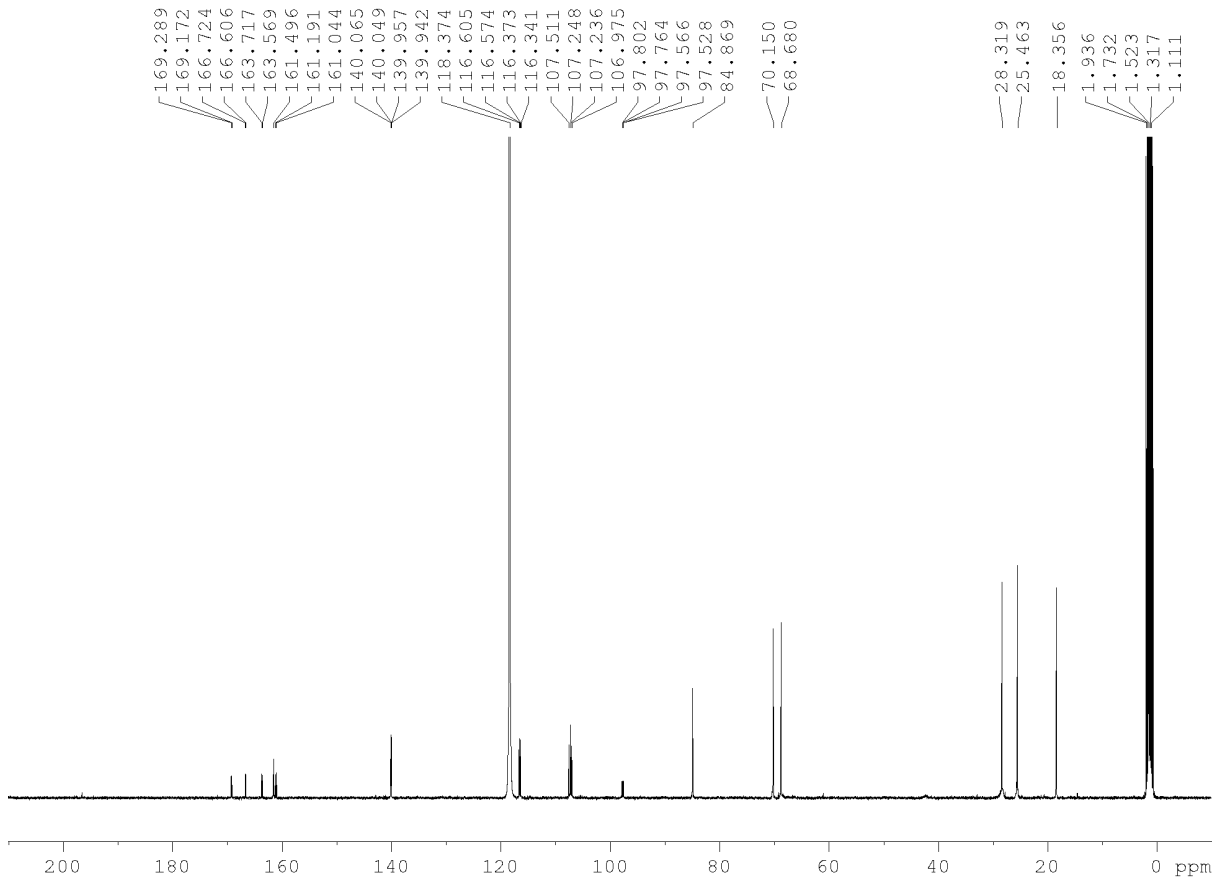
74

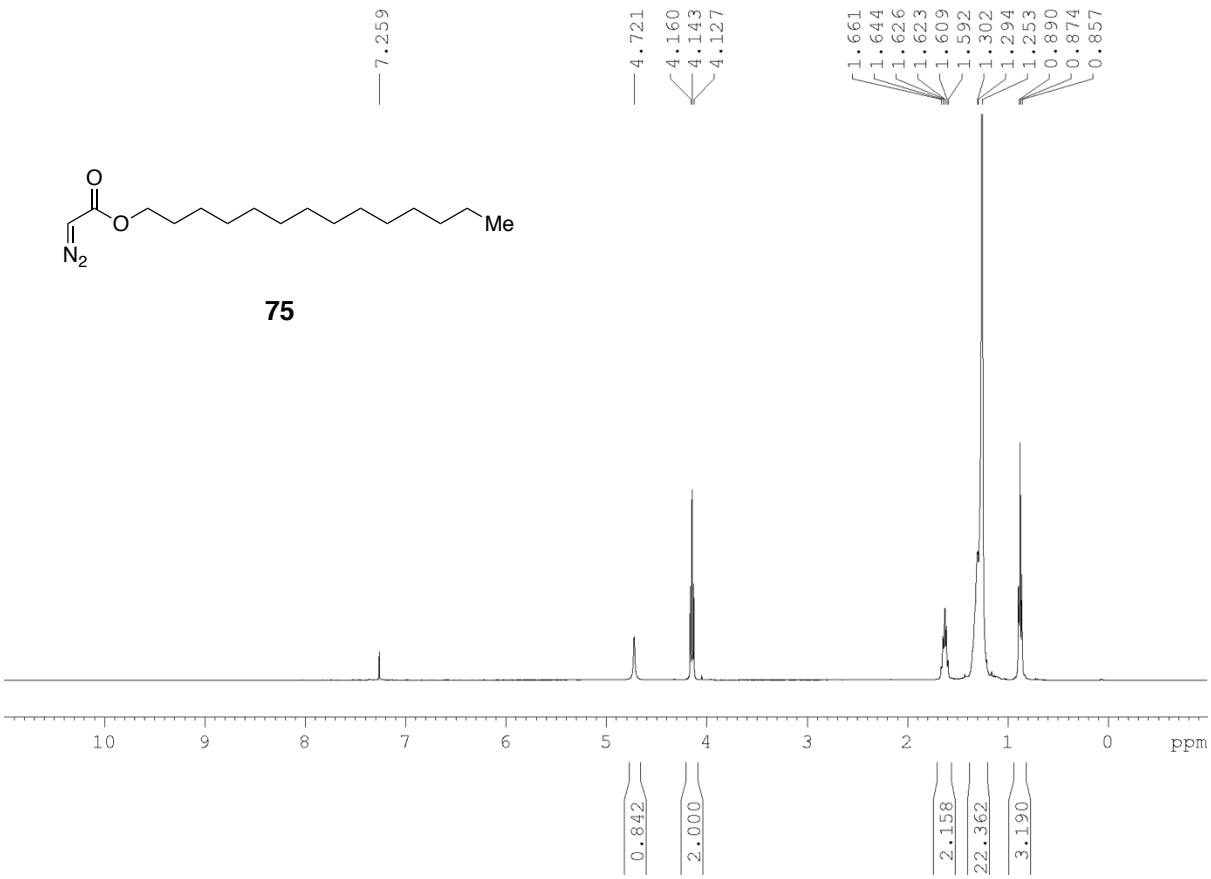
*Impurities



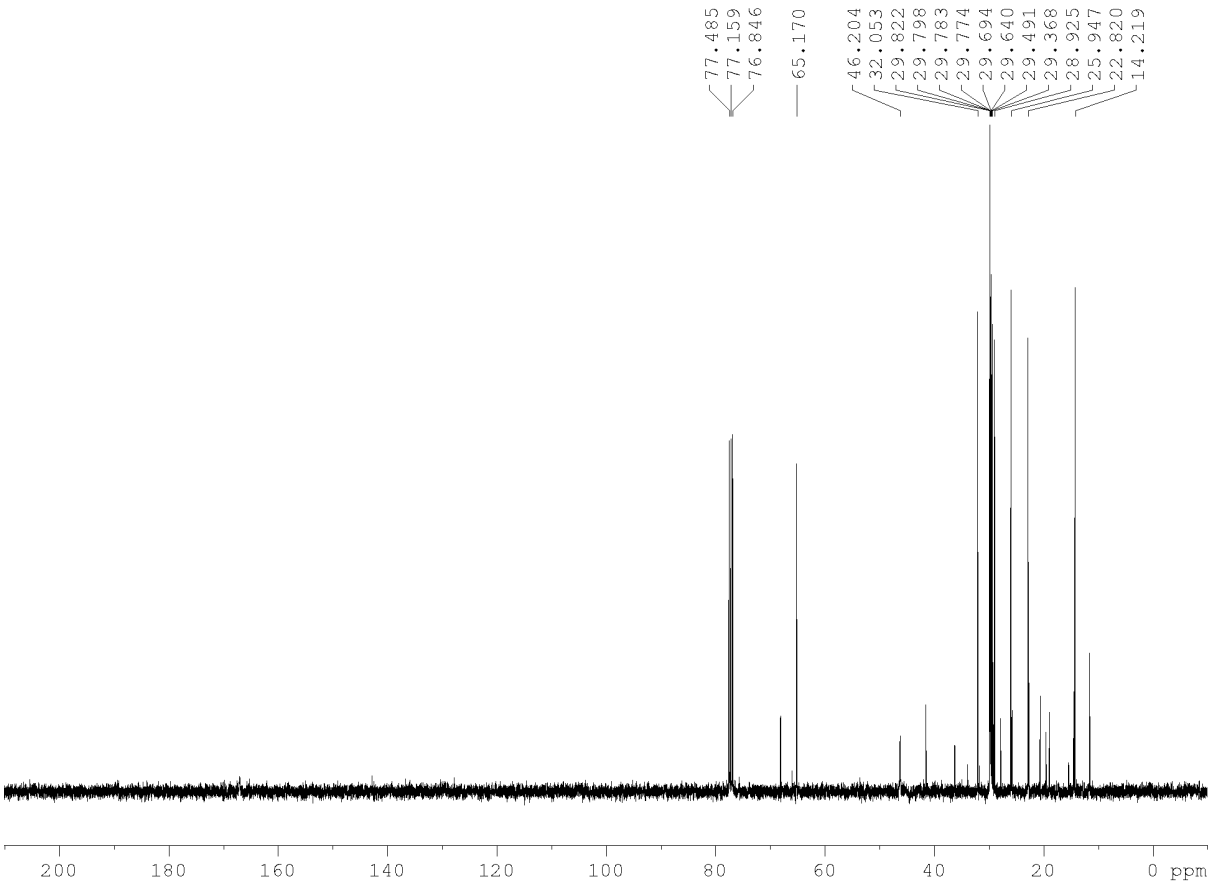

8.244
8.230
8.228
8.222
8.213
8.207
8.205
8.191
7.388
7.381
7.366
7.359
7.344
7.337
7.247
7.244
7.240
7.237
7.227
7.224
7.221
7.220
7.217
7.215
7.204
7.201
7.197
7.195
4.267
4.251
4.236
2.224
2.218
2.207
2.200
2.189
2.185
1.952
1.946
1.940
1.934
1.928
1.779
1.762
1.745
1.739
1.724
1.708
1.559
1.542
1.523
1.504
1.488
F
BF4
F
I
O
N2
O
10
9
8
7
6
5
4
3
2
1
0
ppm
1.000
1.132
1.078
2.322
3.118
2.570
2.410

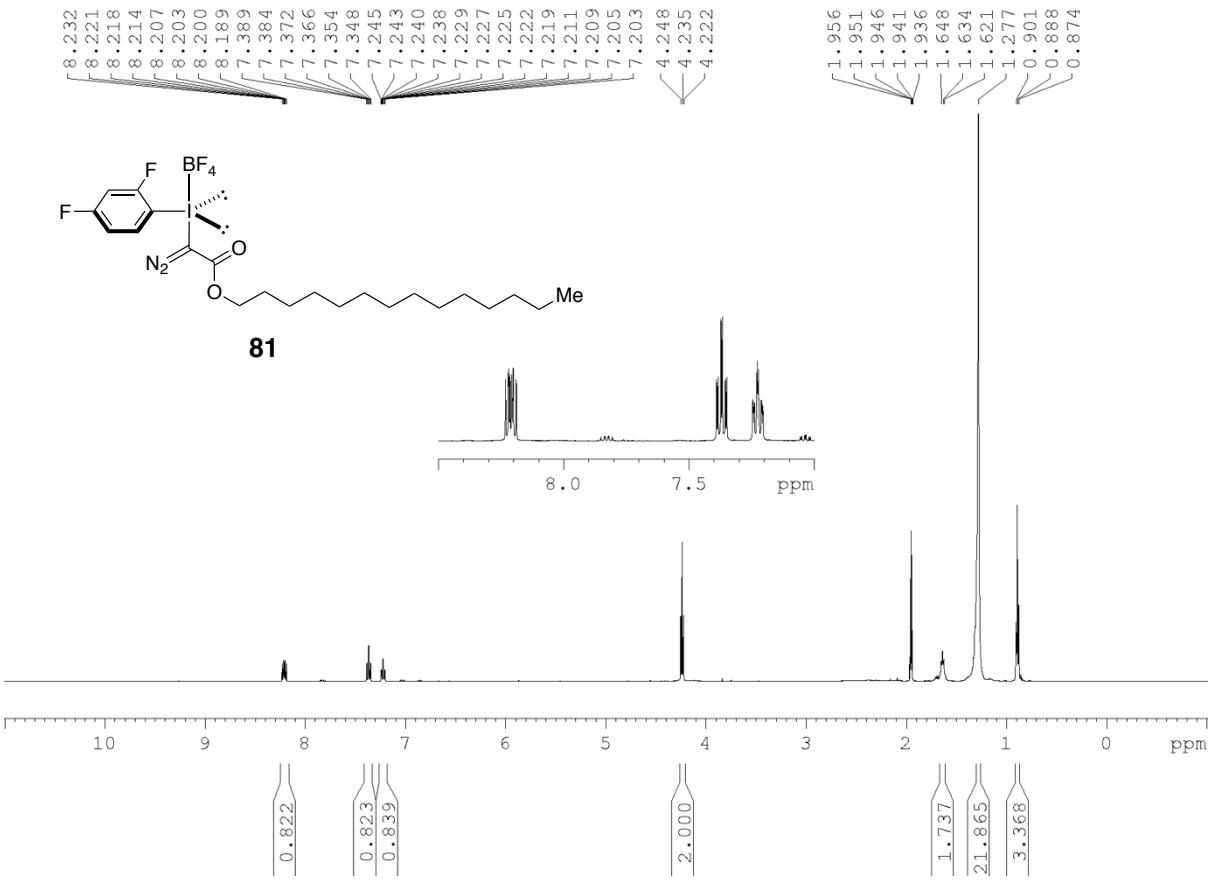
80


169.289
169.172
166.724
166.606
163.717
163.569
161.496
161.191
161.044
140.065
140.049
139.957
139.942
118.374
116.605
116.574
116.373
116.341
107.511
107.248
107.236
106.975
97.802
97.764
97.566
97.528
84.869
70.150
68.680
28.319
25.463
18.356
1.936
1.732
1.523
1.317
1.111
200
180
160
140
120
100
80
60
40
20
0 ppm

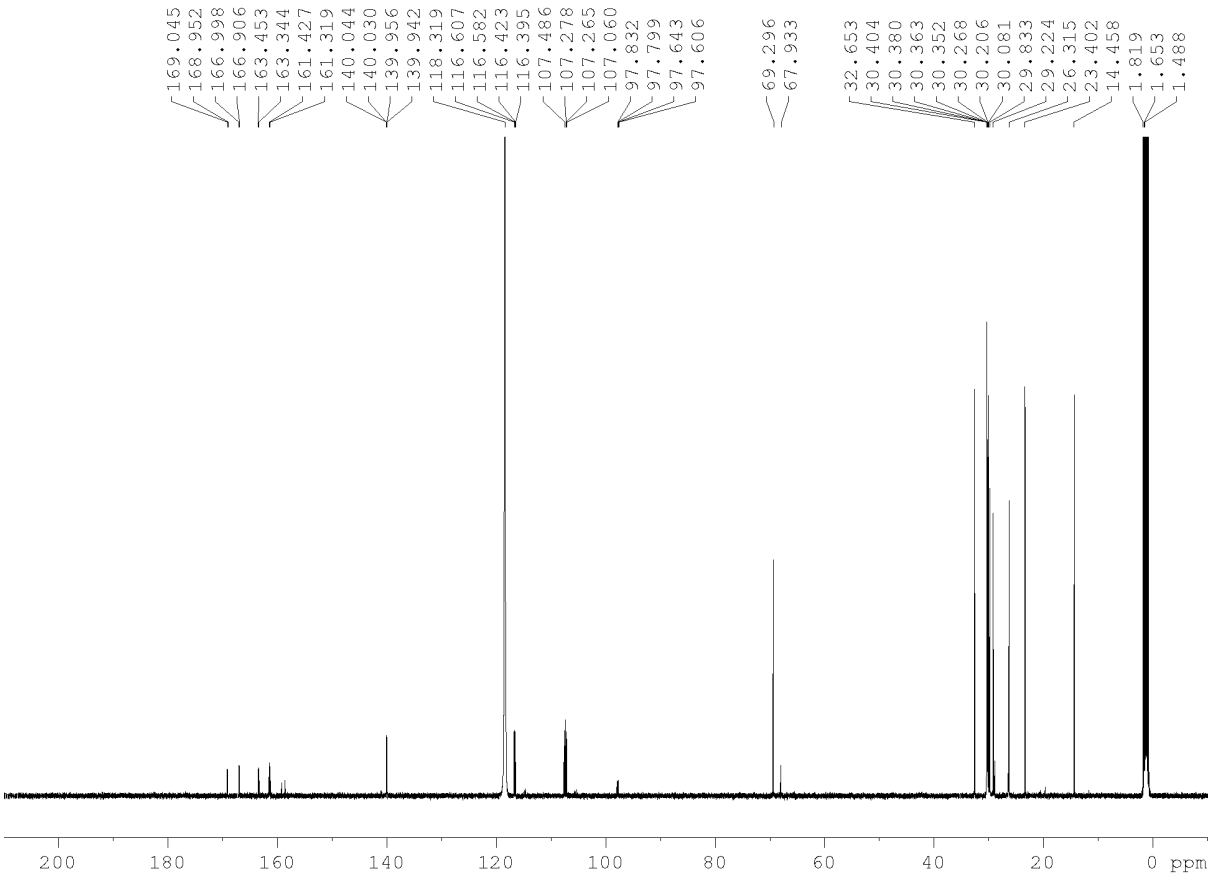

7.259
4.721
4.160
4.143
4.127
1.661
1.644
1.626
1.623
1.609
1.592
1.302
1.294
1.253
0.890
0.874
0.857
O
O
Me
N2
10
9
8
7
6
5
4
3
2
1
0
ppm
0.842
2.000
2.158
22.362
3.190

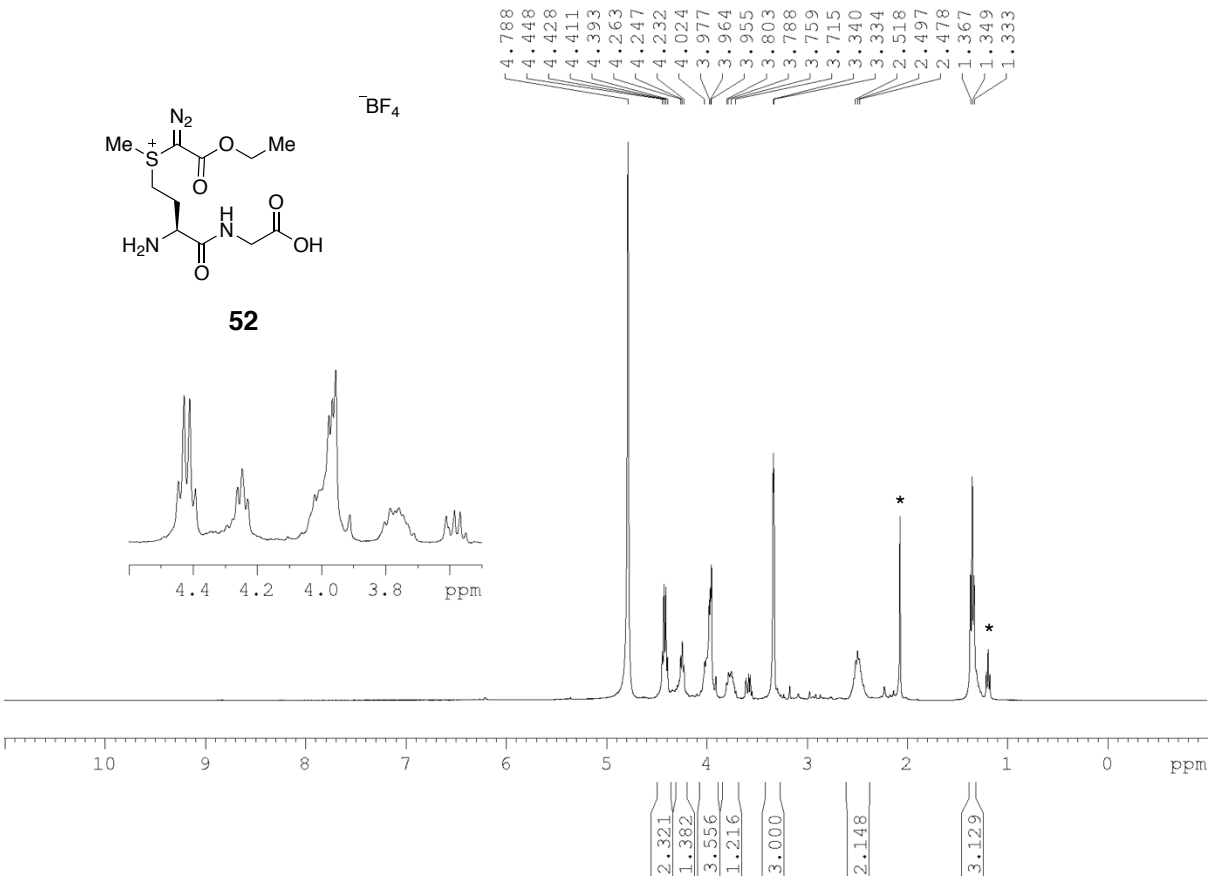
75


77.485
77.159
76.846
65.170
46.204
32.053
29.822
29.798
29.783
29.774
29.694
29.640
29.491
29.368
28.925
25.947
22.820
14.219
200
180
160
140
120
100
80
60
40
20
0 ppm

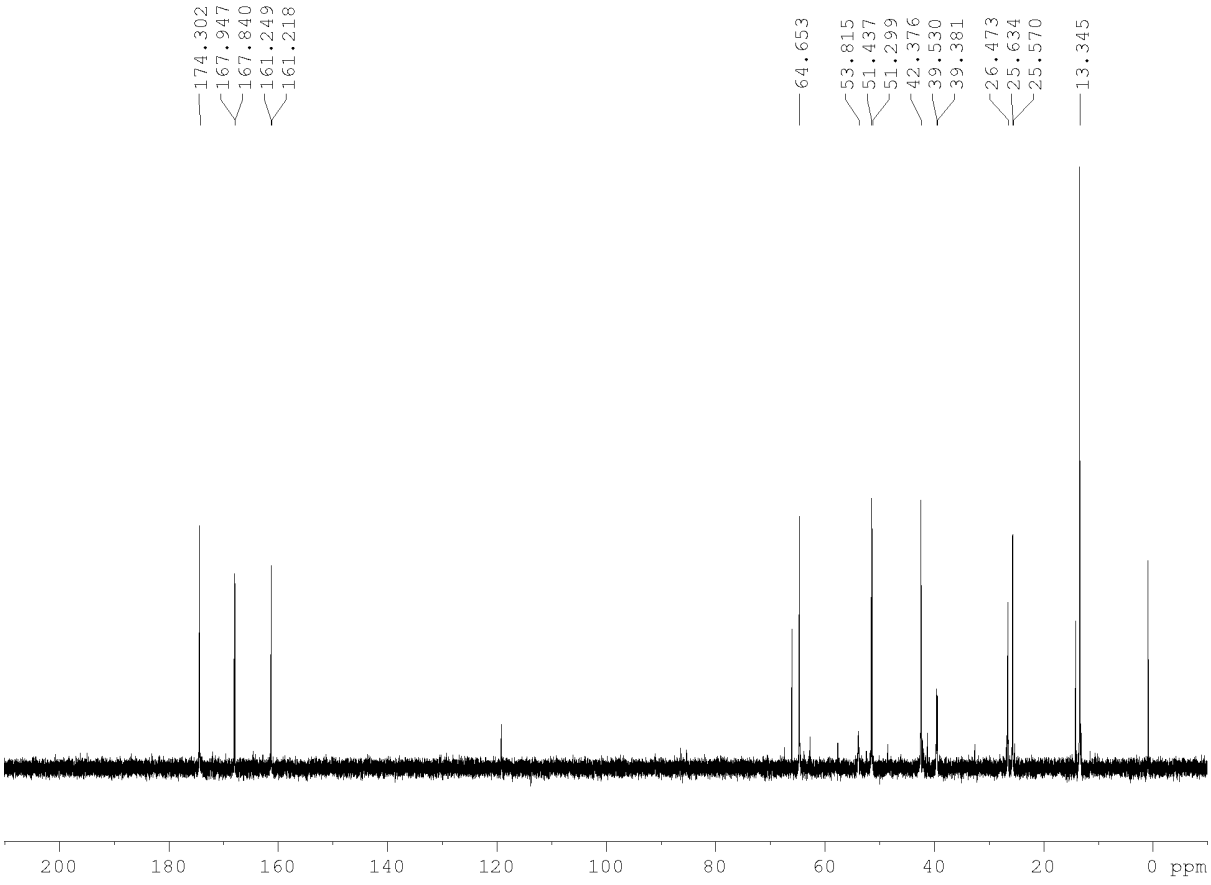

8.232
8.221
8.218
8.214
8.207
8.203
8.200
8.189
7.389
7.384
7.372
7.366
7.354
7.348
7.245
7.243
7.240
7.238
7.229
7.227
7.225
7.222
7.219
7.211
7.209
7.205
7.203
4.248
4.235
4.222
1.956
1.951
1.946
1.941
1.936
1.648
1.634
1.621
1.277
0.901
0.888
0.874
F
BF4
F
I
O
N2
O
Me
8.0
7.5
ppm
10
9
8
7
6
5
4
3
2
1
0
ppm
0.822
0.823
0.839
2.000
1.737
21.865
3.368

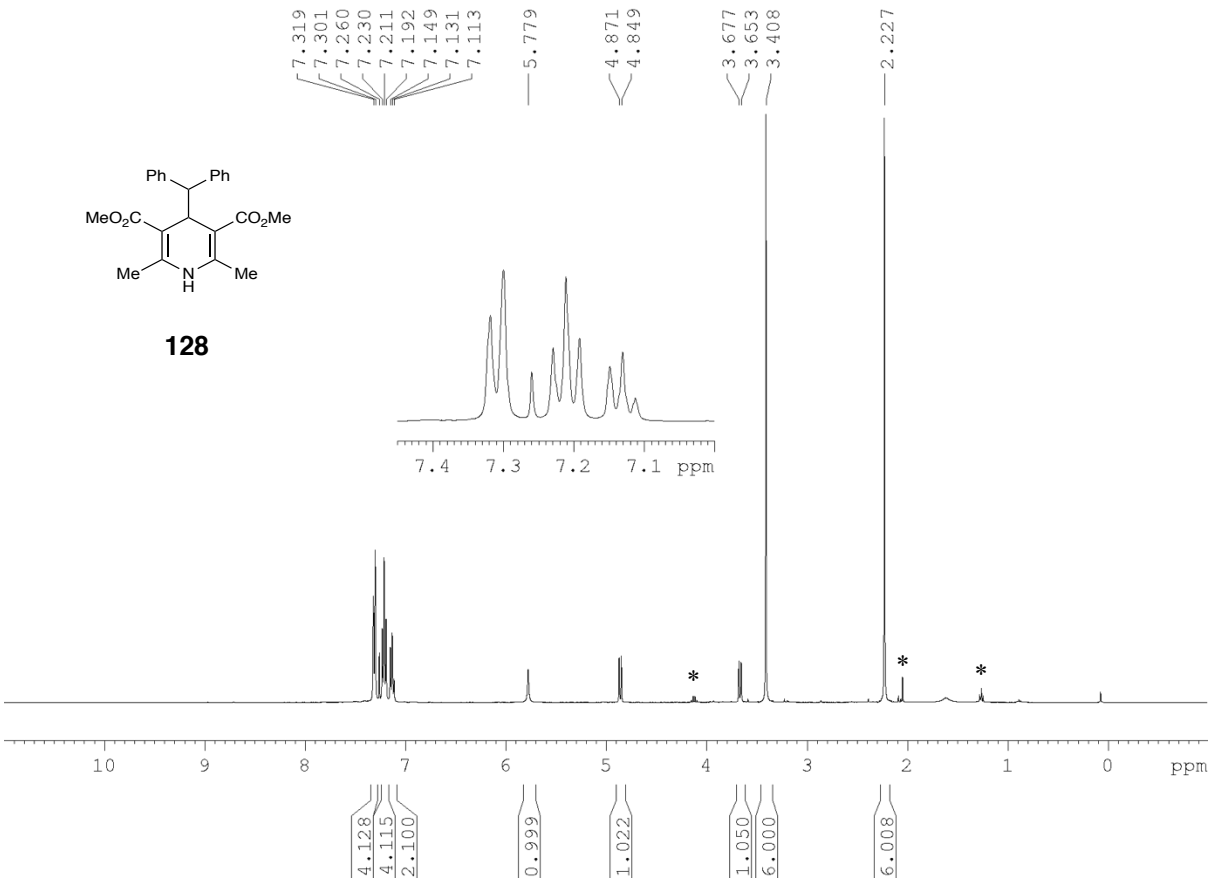
81


169.045
168.952
166.998
166.906
163.453
163.344
161.427
161.319
140.044
140.030
139.956
139.942
118.319
116.607
116.582
116.423
116.395
107.486
107.278
107.265
107.060
97.832
97.799
97.643
97.606
69.296
67.933
32.653
30.404
30.380
30.363
30.352
30.268
30.206
30.081
29.833
29.224
26.315
23.402
14.458
1.819
1.653
1.488
200
180
160
140
120
100
80
60
40
20
0 ppm


4.788
4.448
4.428
4.411
4.393
4.263
4.247
4.232
4.024
3.977
3.964
3.955
3.803
3.788
3.759
3.715
3.340
3.334
2.518
2.497
2.478
1.367
1.349
1.333
$^{-}BF_4$
Me
S
N2
O
Me
H2N
H
N
OH
52
4.4
4.2
4.0
3.8
ppm
*
*
10
9
8
7
6
5
4
3
2
1
0
ppm
2.321
1.382
3.556
1.216
3.000
2.148
3.129

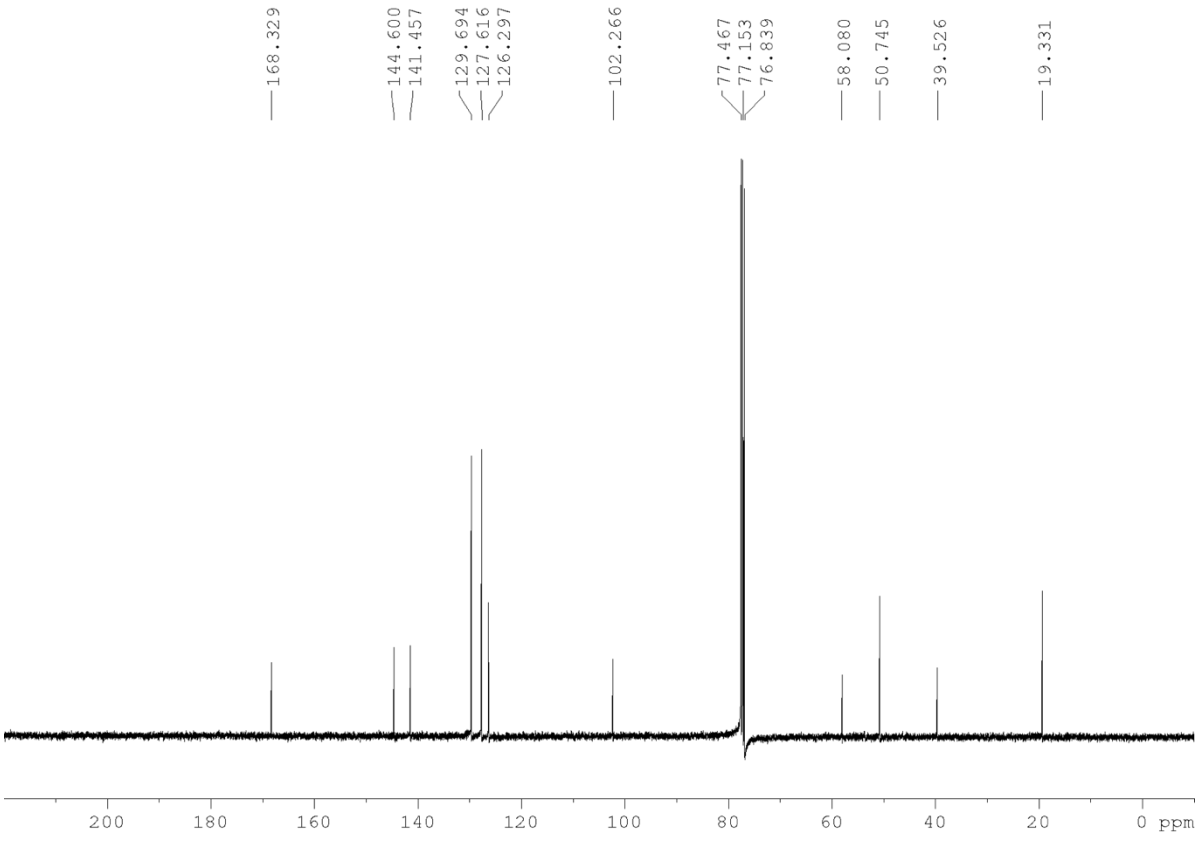
*Impurities

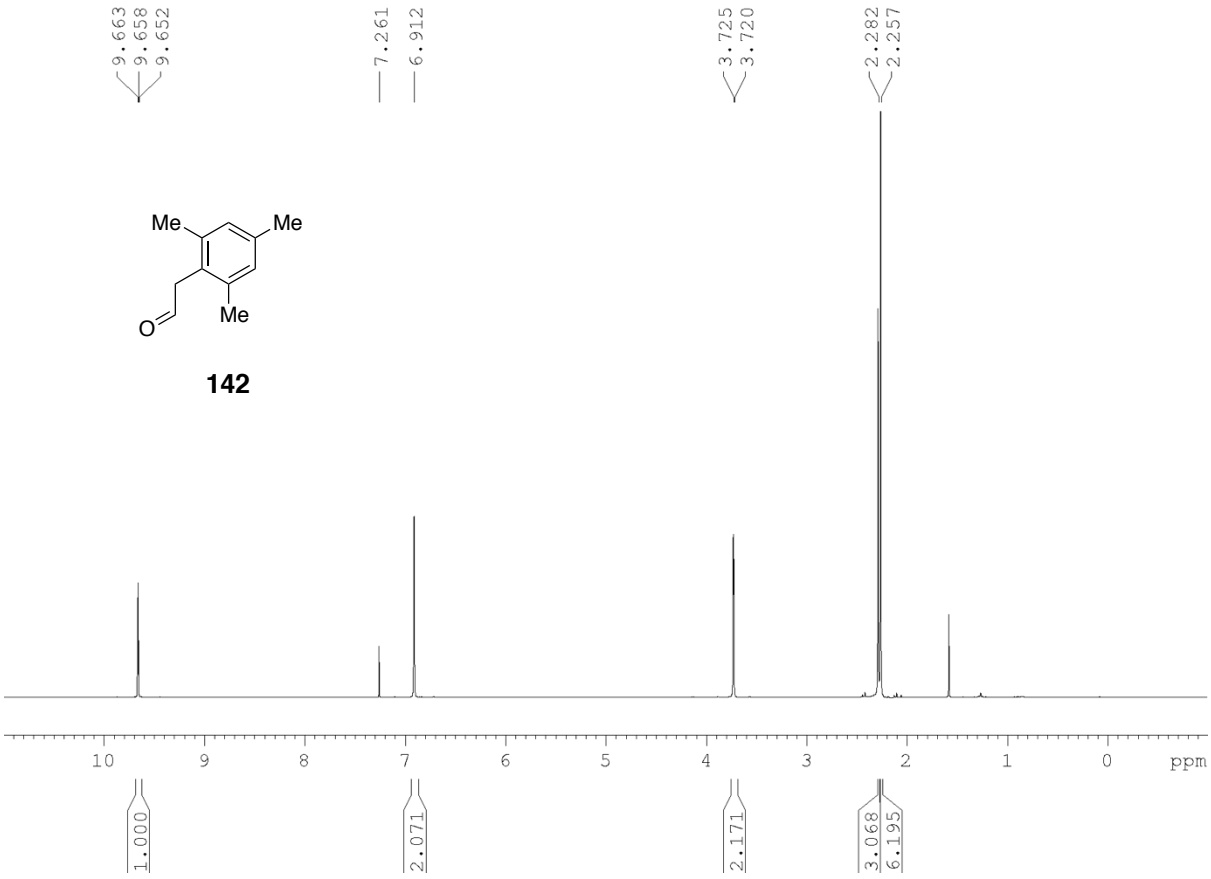

174.302
167.947
167.840
161.249
161.218
64.653
53.815
51.437
51.299
42.376
39.530
39.381
26.473
25.634
25.570
13.345
200
180
160
140
120
100
80
60
40
20
0 ppm



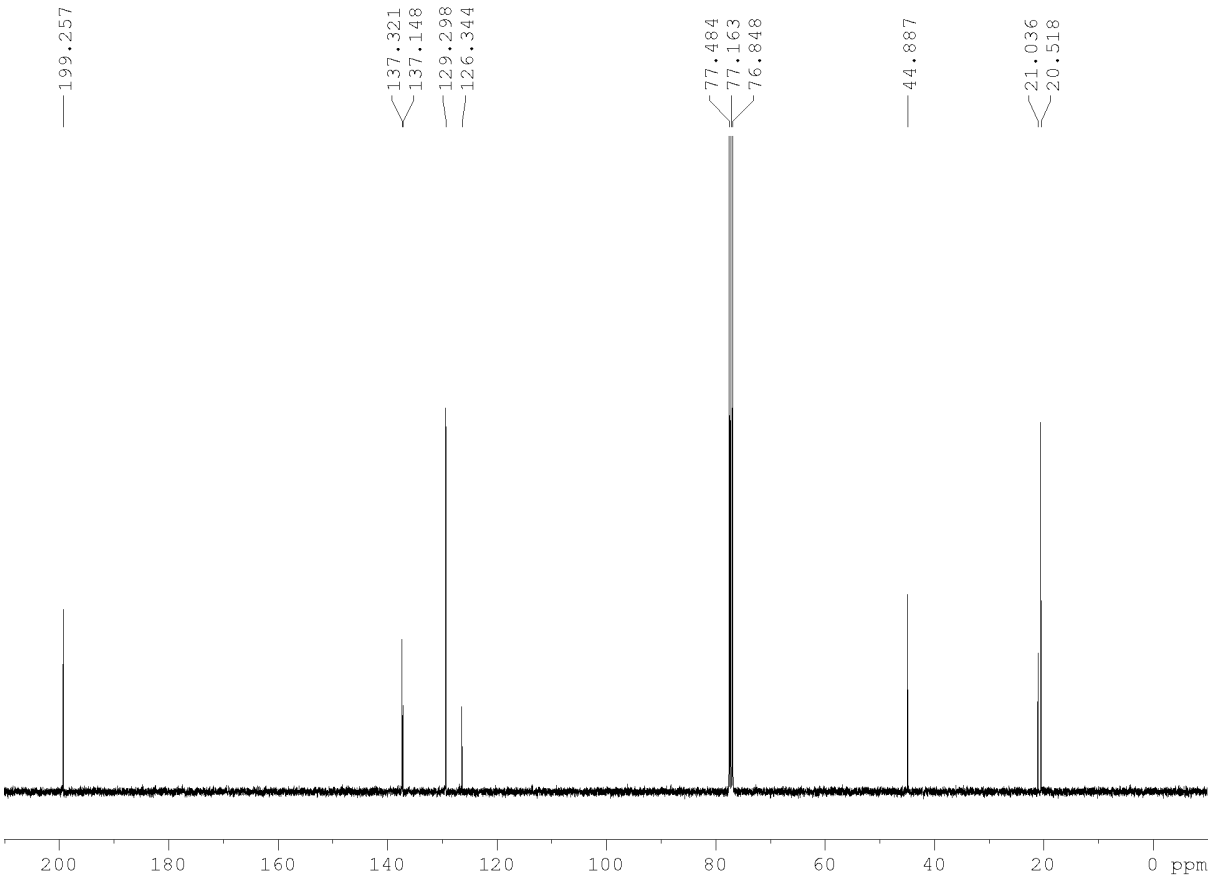
128

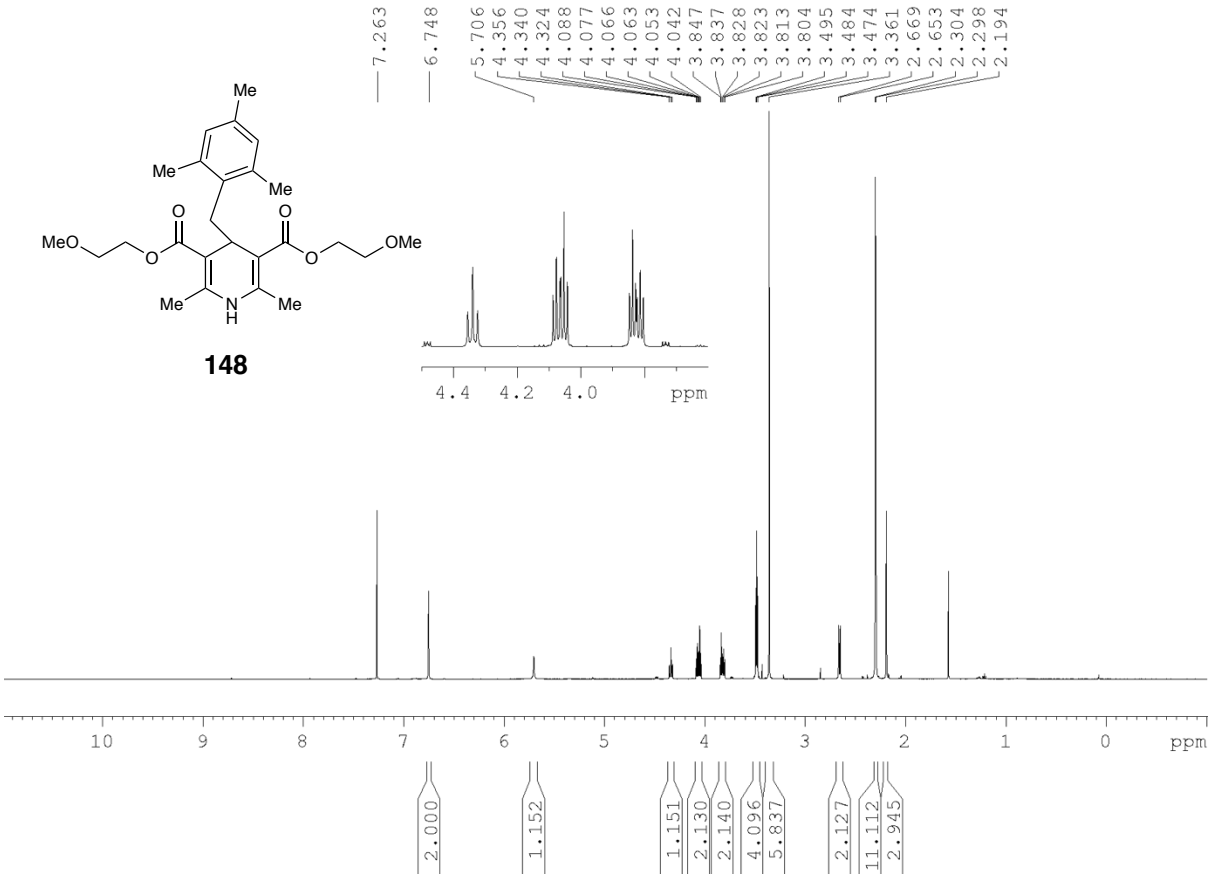
* Impurities



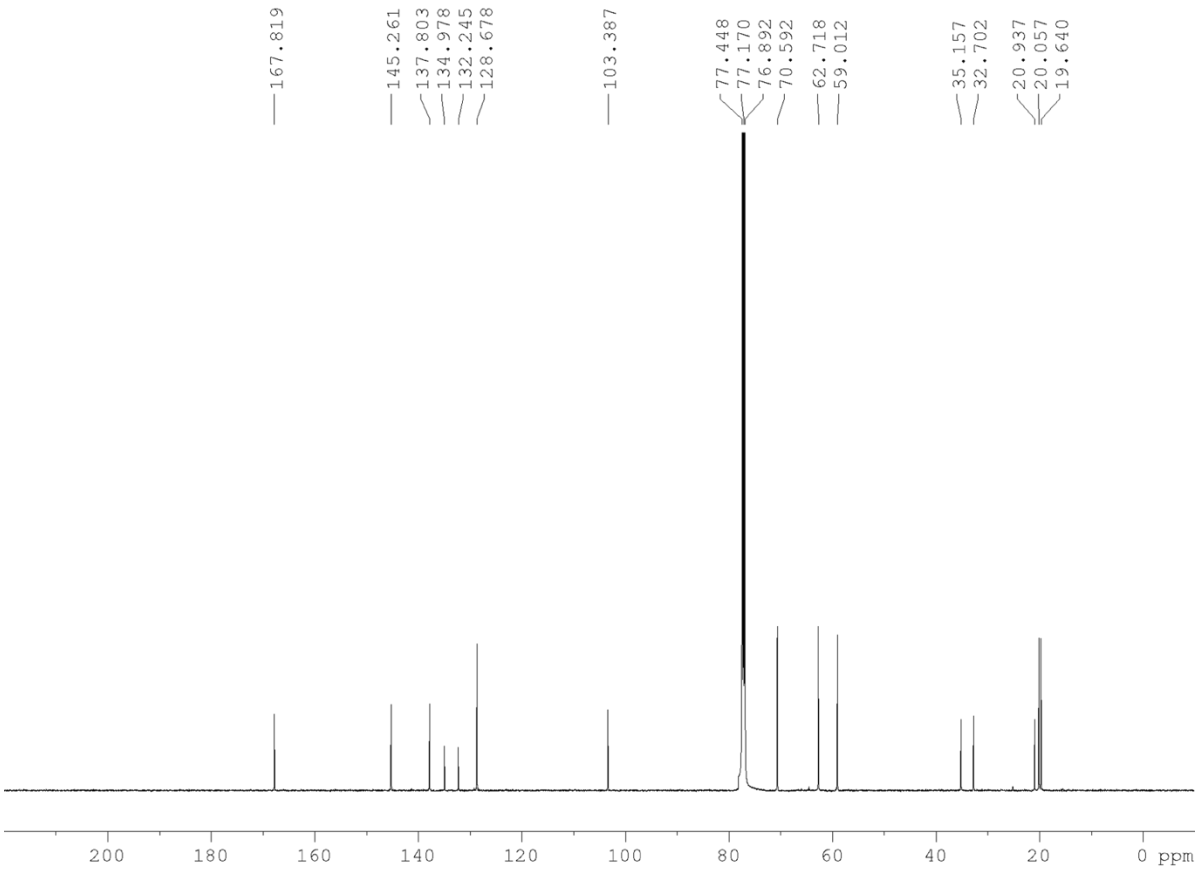

9.663
9.658
9.652
7.261
6.912
3.725
3.720
2.282
2.257
Me
Me
Me
O
10
9
8
7
6
5
4
3
2
1
0
ppm
1.000
2.071
2.171
3.068
6.195

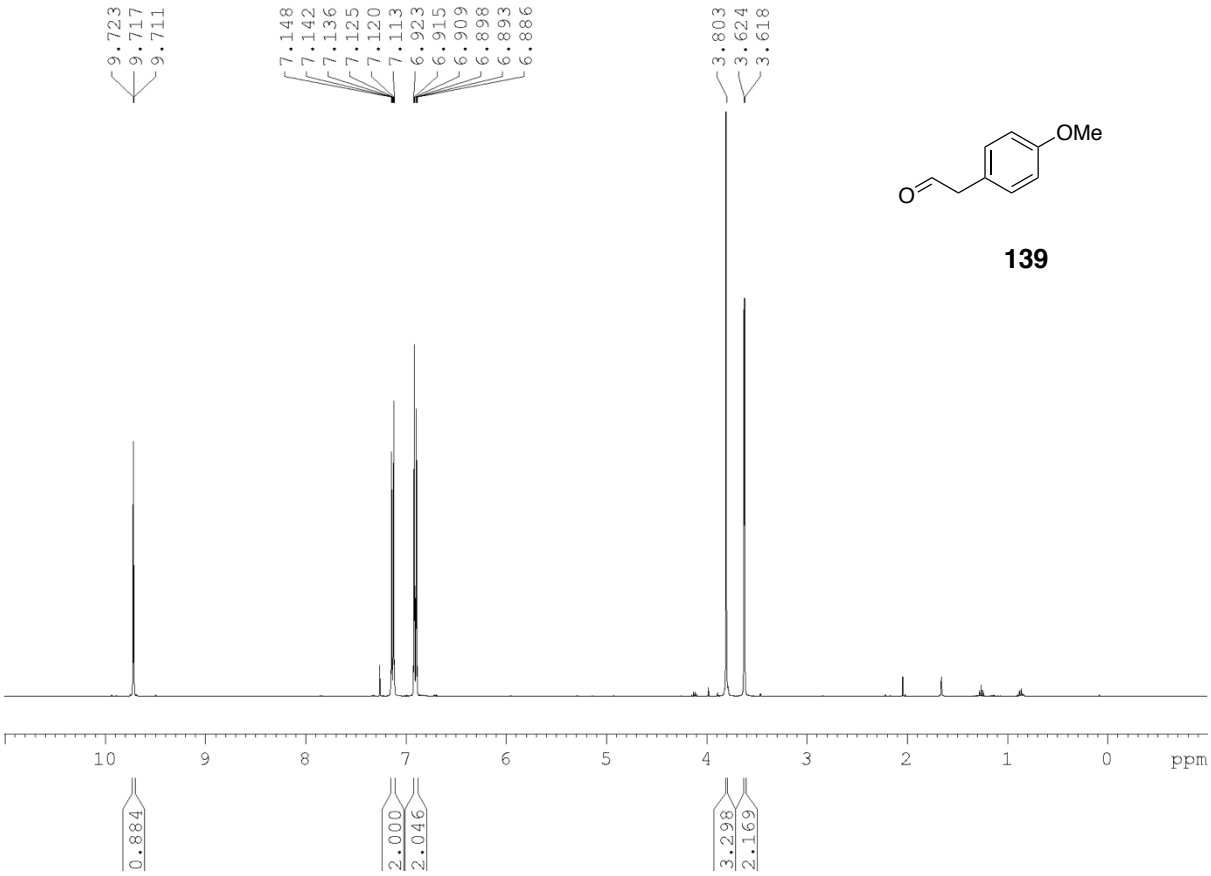
142


199.257
137.321
137.148
129.298
126.344
77.484
77.163
76.848
44.887
21.036
20.518
200
180
160
140
120
100
80
60
40
20
0 ppm

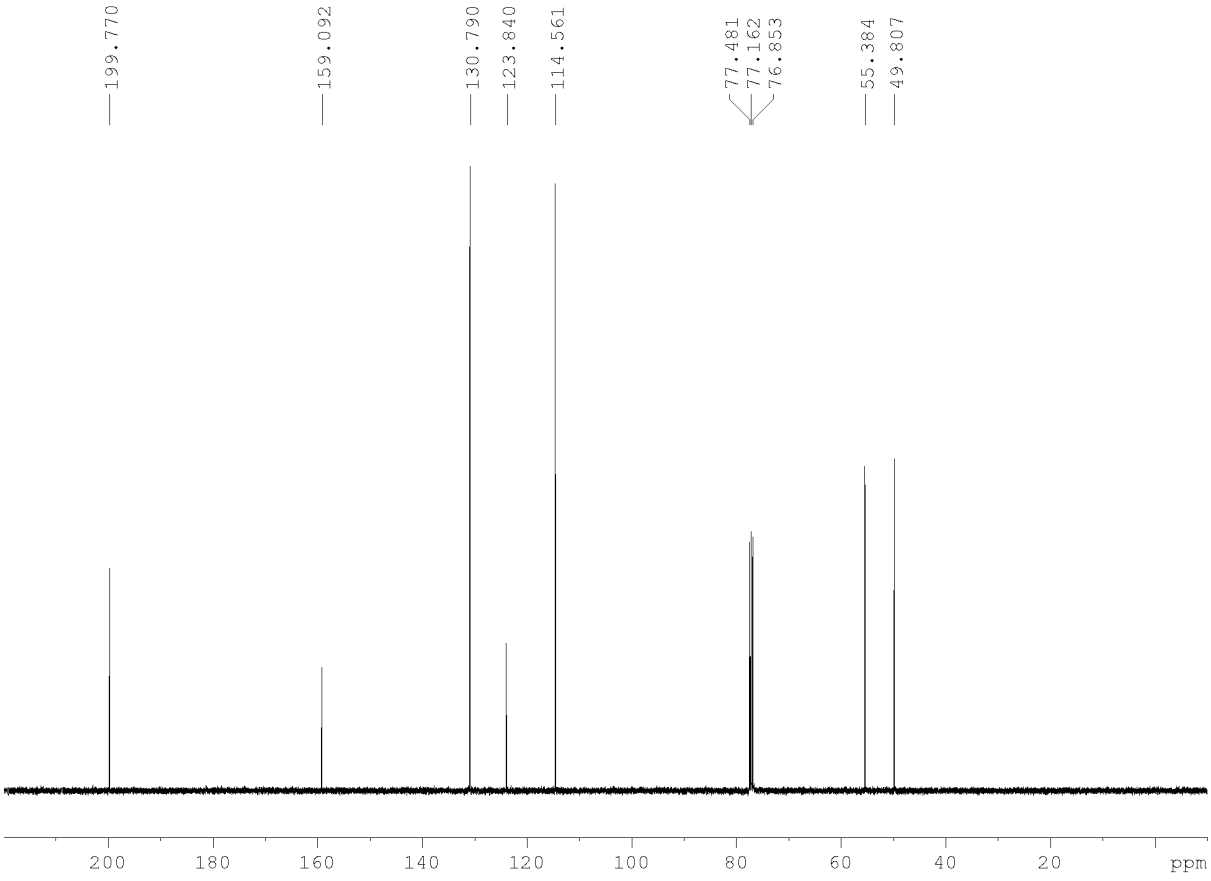

7.263
6.748
5.706
4.356
4.340
4.324
4.088
4.077
4.066
4.063
4.053
4.042
3.847
3.837
3.828
3.823
3.813
3.804
3.495
3.484
3.474
3.361
2.669
2.653
2.304
2.298
2.194
Me
Me
Me
O
O
MeO
O
O
OMe
Me
N
H
Me
4.4
4.2
4.0
ppm
10
9
8
7
6
5
4
3
2
1
0
ppm
2.000
1.152
1.151
2.130
2.140
4.096
5.837
2.127
11.112
2.945

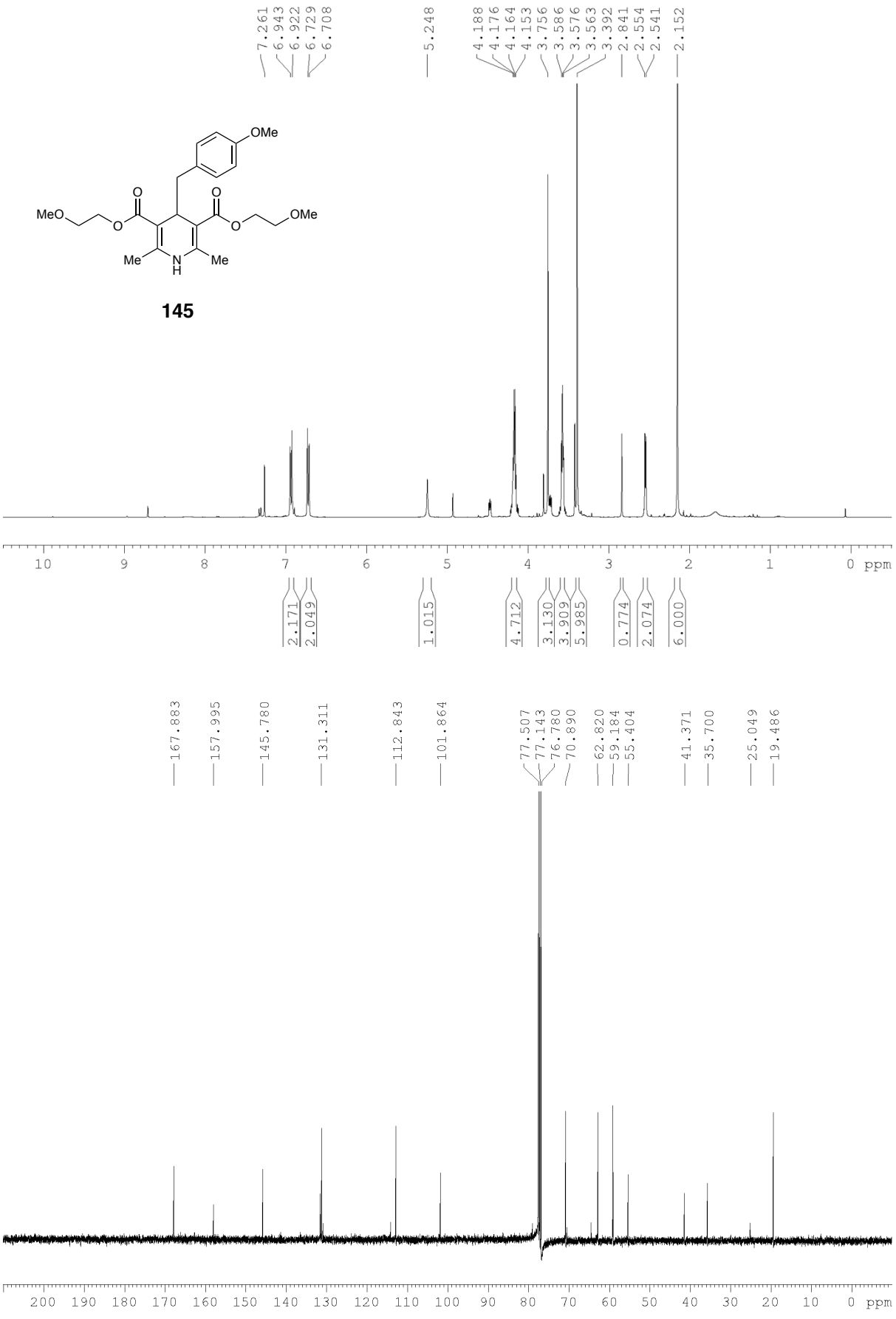
148

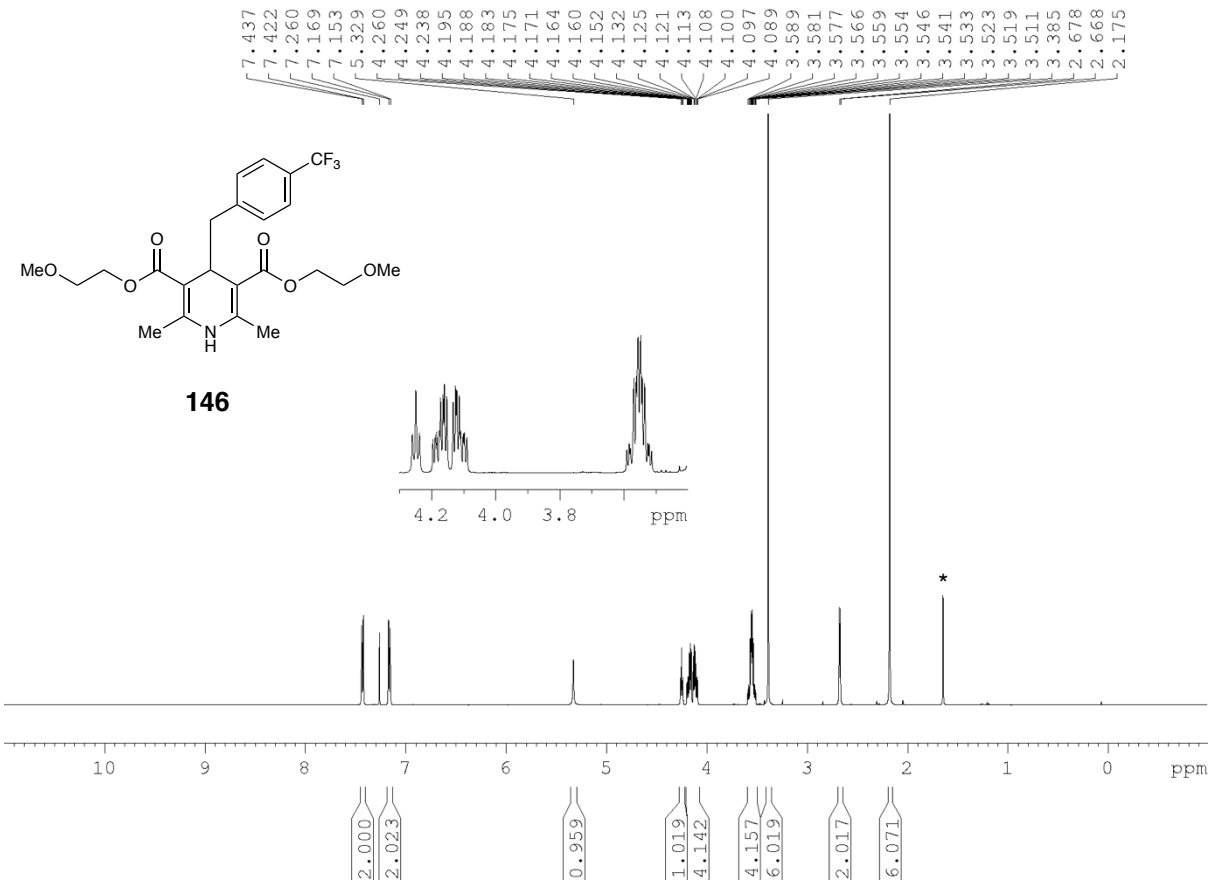

167.819
145.261
137.803
134.978
132.245
128.678
103.387
77.448
77.170
76.892
70.592
62.718
59.012
35.157
32.702
20.937
20.057
19.640
200
180
160
140
120
100
80
60
40
20
0 ppm


9.723
9.717
9.711
7.148
7.142
7.136
7.125
7.120
7.113
6.923
6.915
6.909
6.898
6.893
6.886
3.803
3.624
3.618
OMe
O
10
9
8
7
6
5
4
3
2
1
0
ppm
0.884
2.000
2.046
3.298
2.169

139

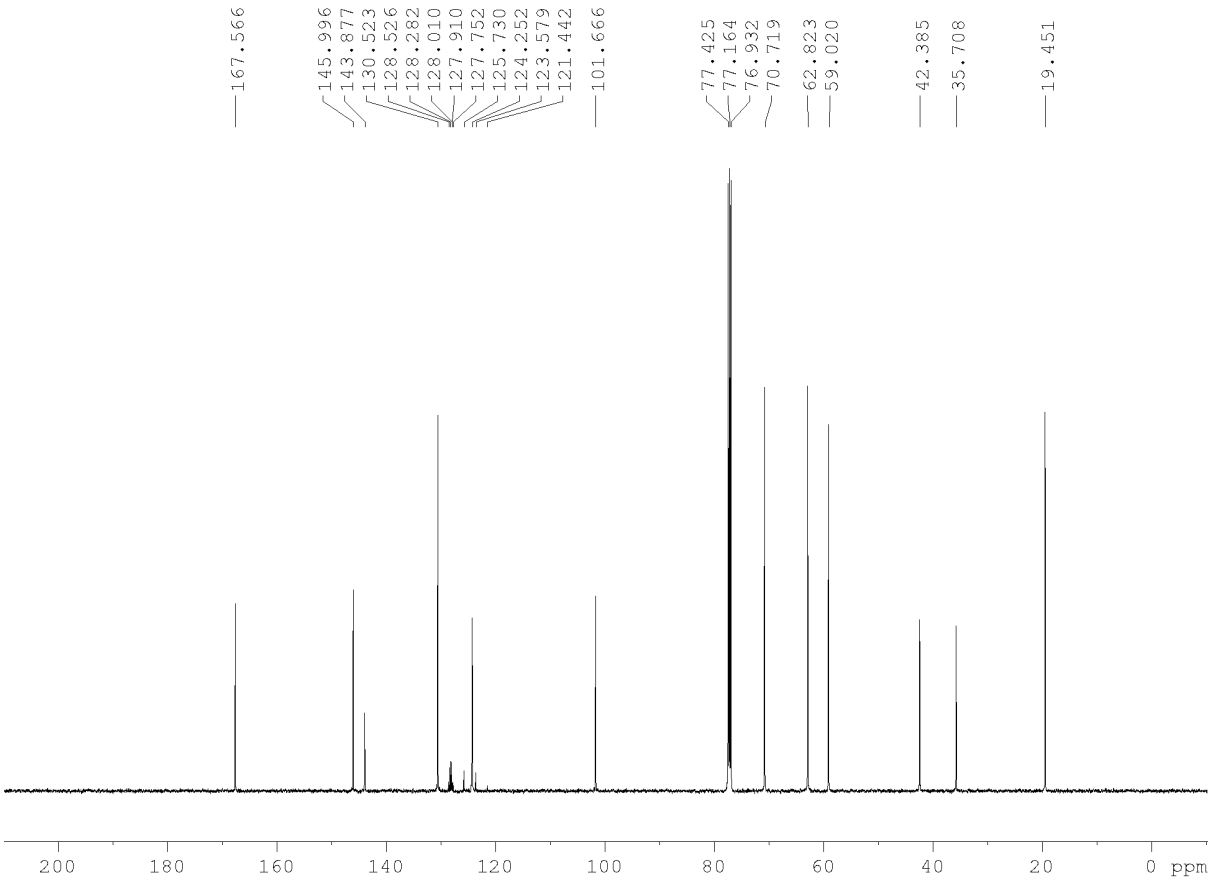

199.770
159.092
130.790
123.840
114.561
77.481
77.162
76.853
55.384
49.807
200
180
160
140
120
100
80
60
40
20
ppm

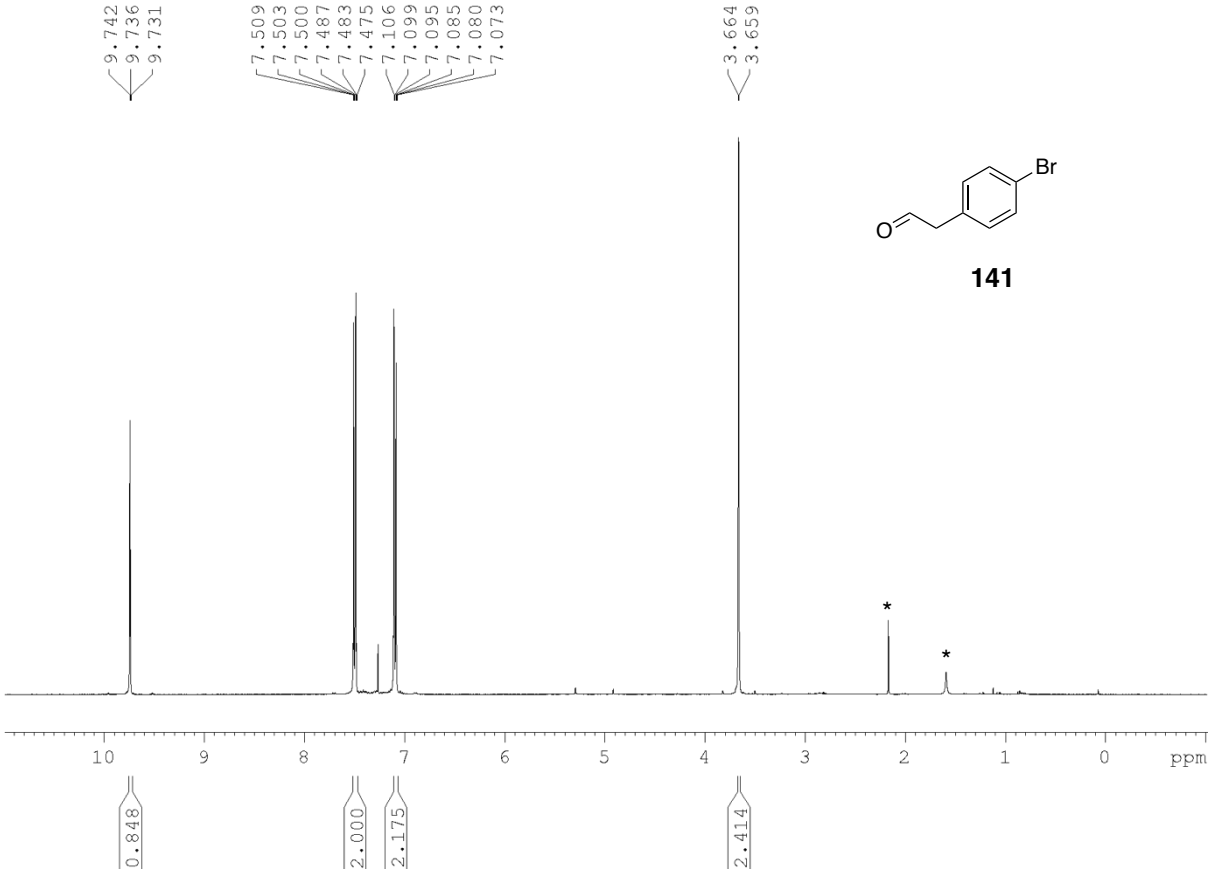

7.261
6.943
6.922
6.729
6.708
5.248
4.188
4.176
4.164
4.153
3.756
3.586
3.576
3.563
3.392
2.841
2.554
2.541
2.152
OMe
MeO
O
O
O
O
OMe
Me
N
H
Me
145
10 9 8 7 6 5 4 3 2 1 0 ppm
2.171
2.049
1.015
4.712
3.130
3.909
5.985
0.774
2.074
6.000
167.883
157.995
145.780
131.311
112.843
101.864
77.507
77.143
76.780
70.890
62.820
59.184
55.404
41.371
35.700
25.049
19.486
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 ppm



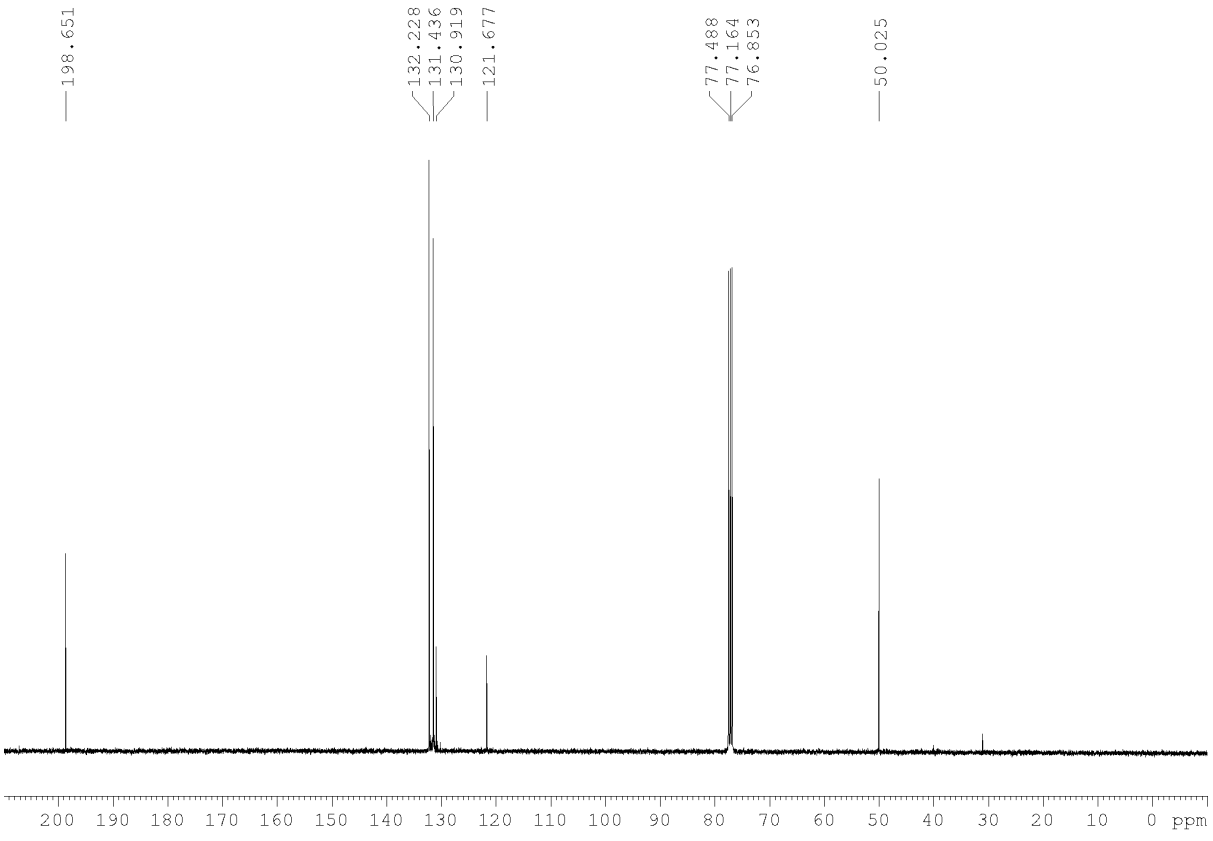
146

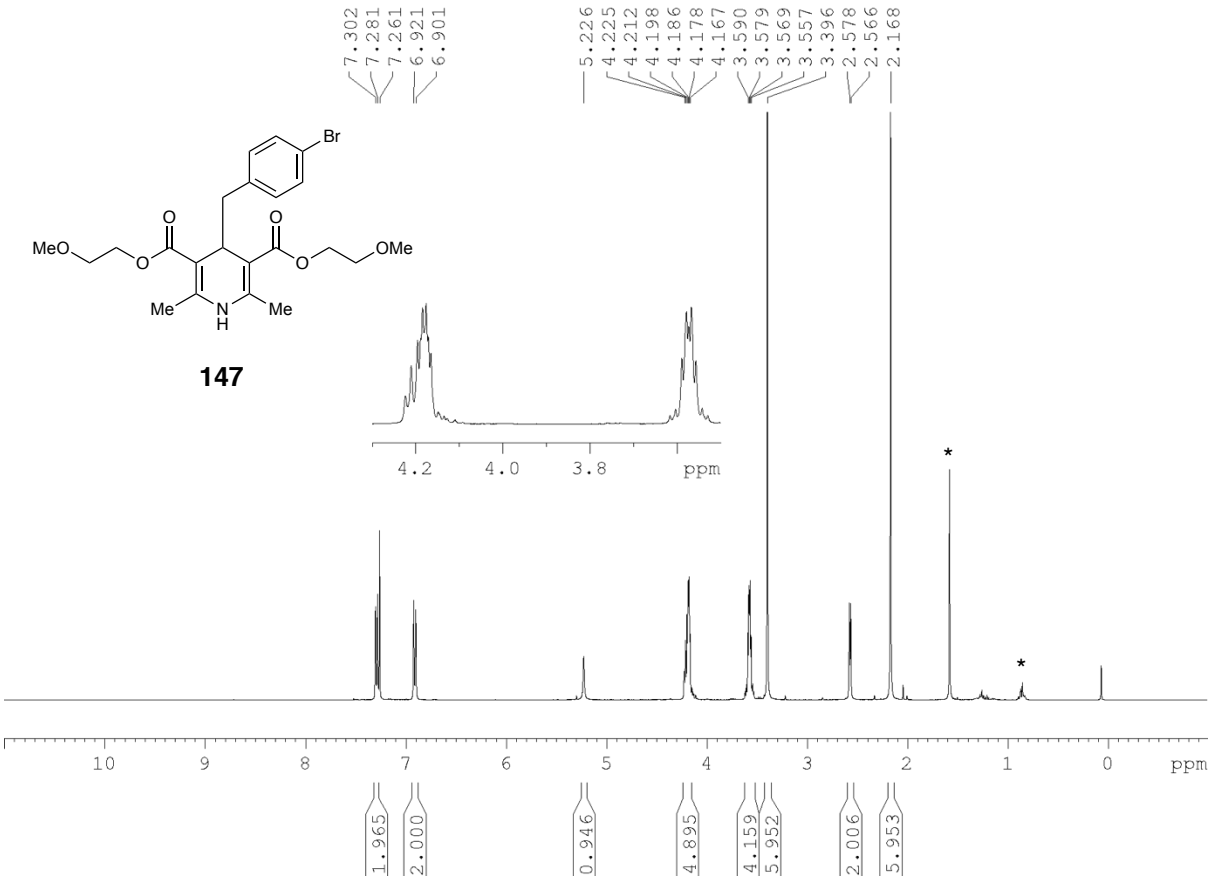
*Impurities




9.742
9.736
9.731
7.509
7.503
7.500
7.487
7.483
7.475
7.106
7.099
7.095
7.085
7.080
7.073
3.664
3.659
Br
O
141
*
*
10
9
8
7
6
5
4
3
2
1
0
ppm
0.848
2.000
2.175
2.414

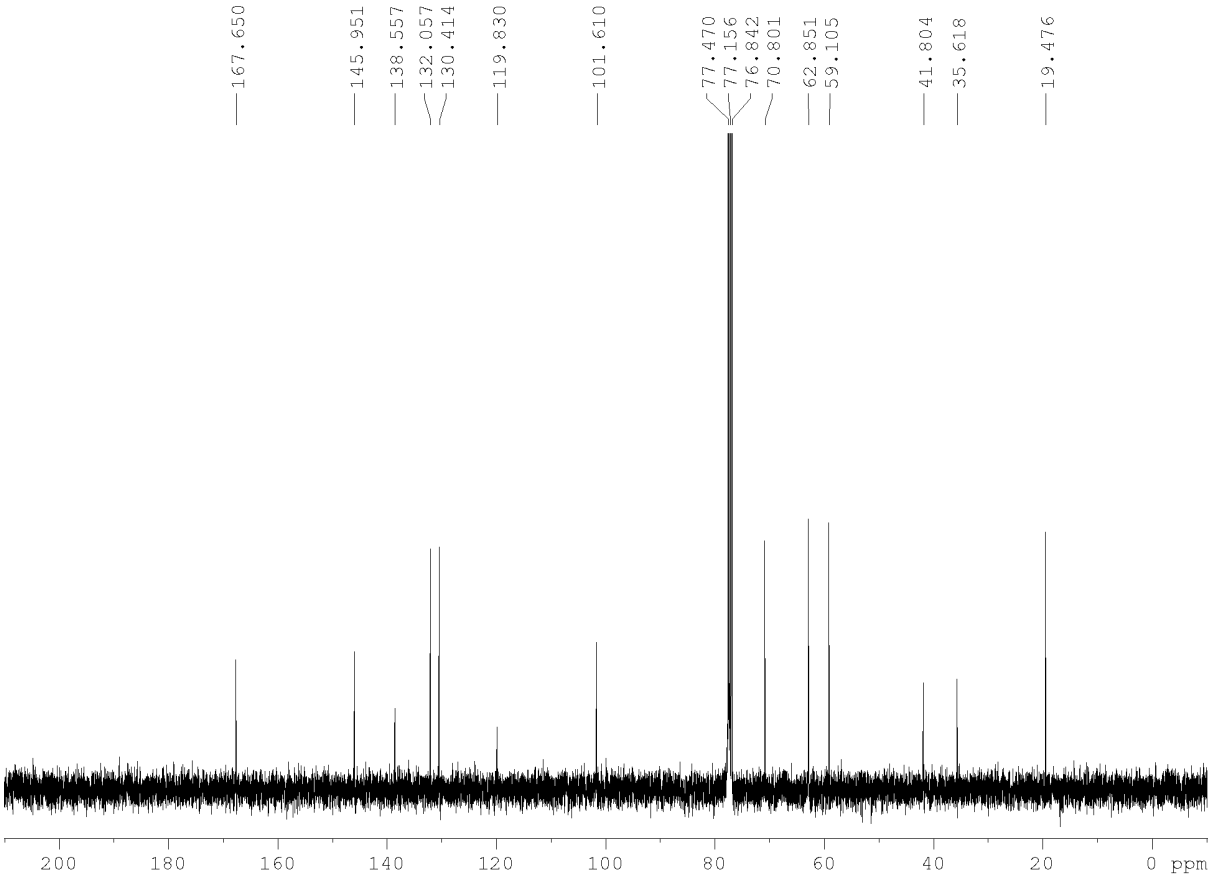
*Impurities

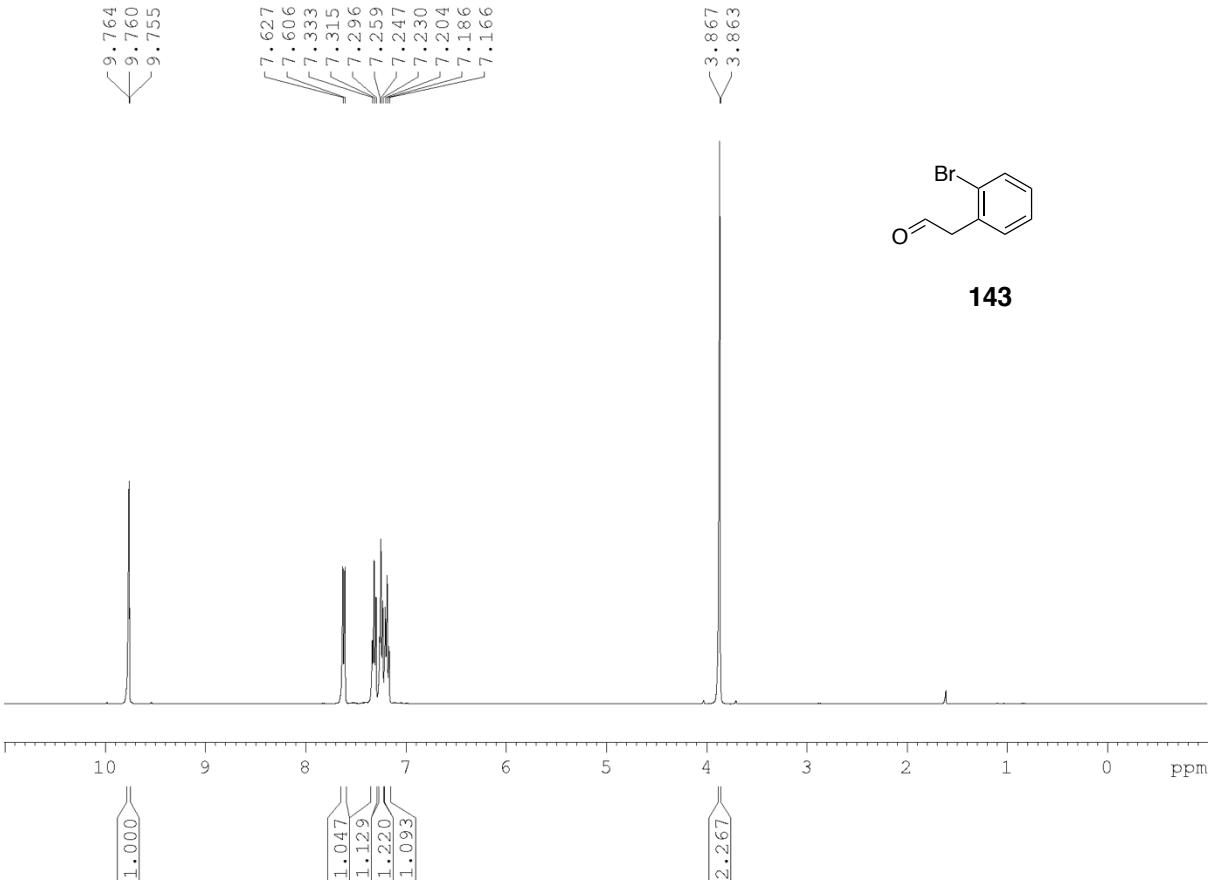

198.651
132.228
131.436
130.919
121.677
77.488
77.164
76.853
50.025
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
ppm

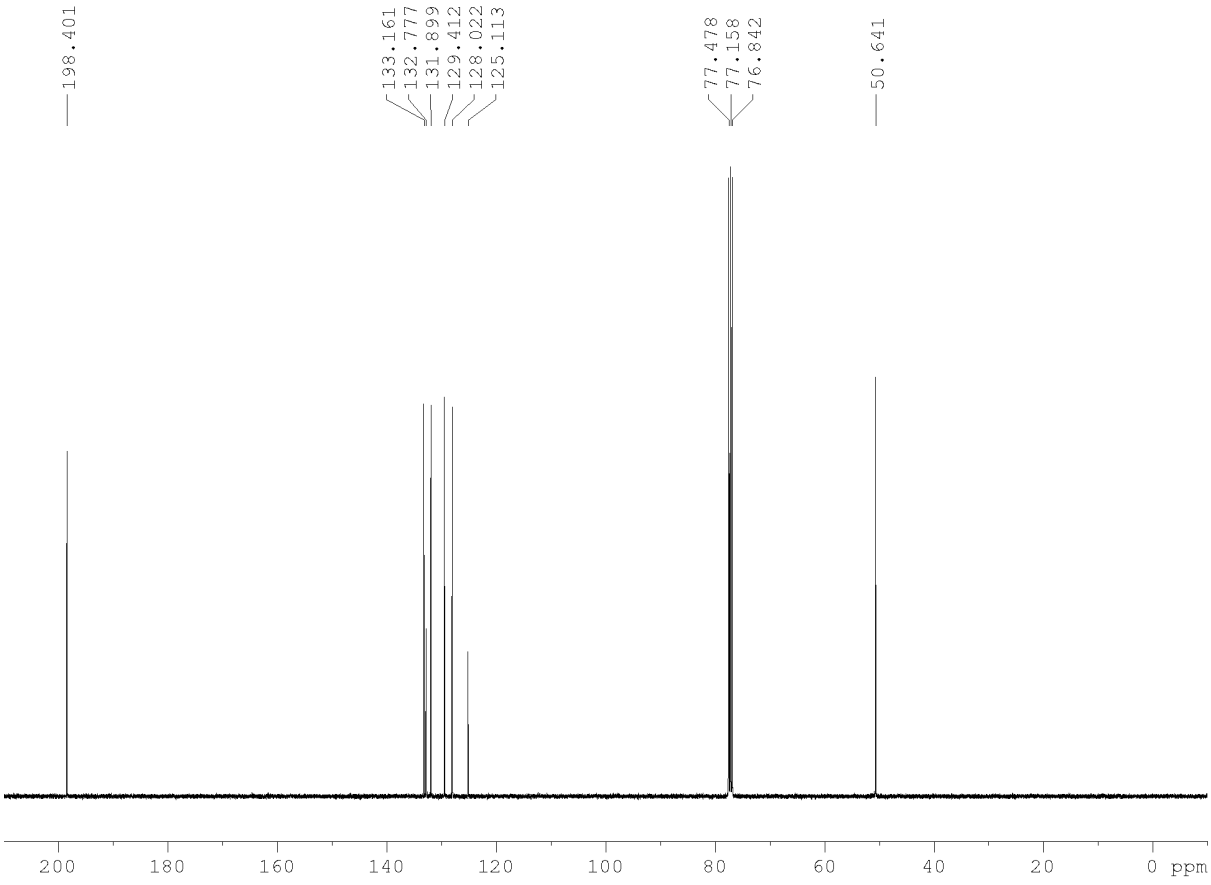


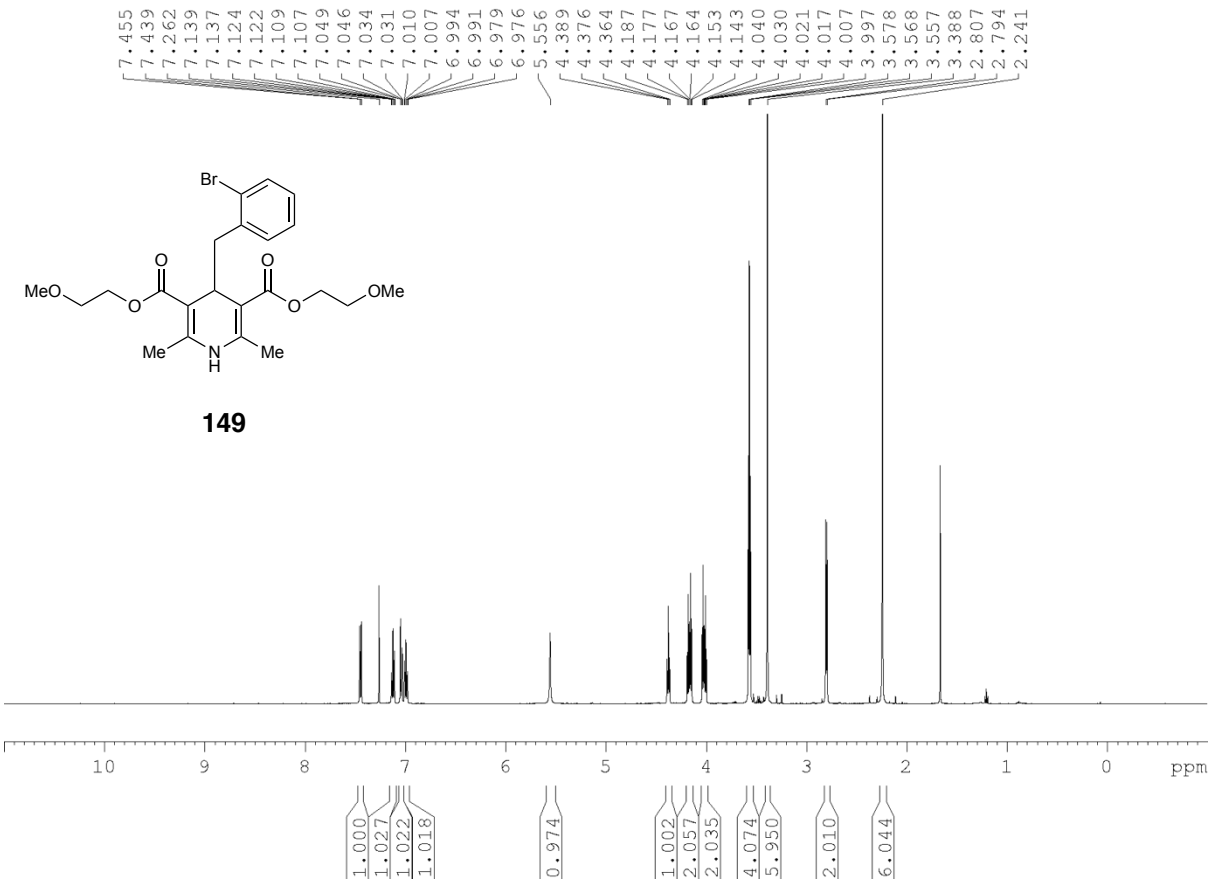
147

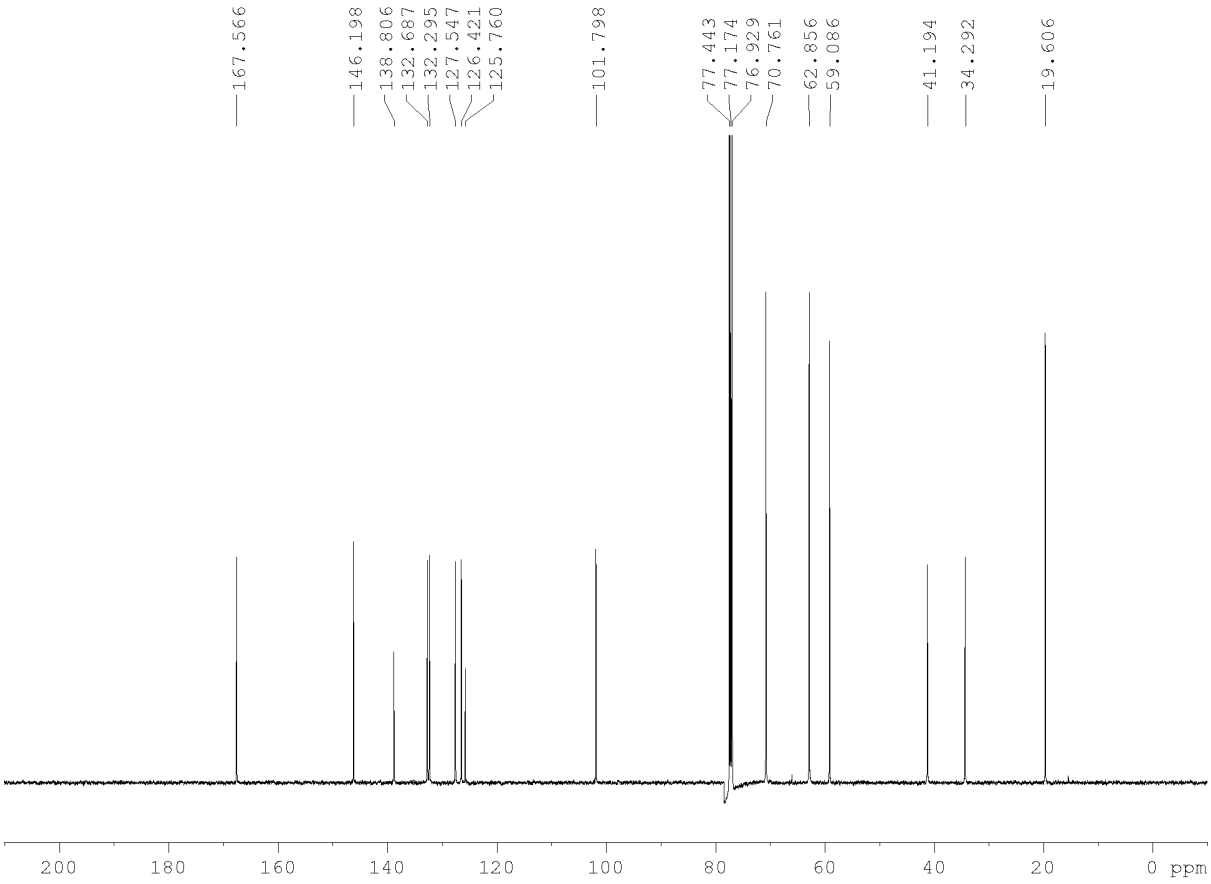
*Impurities

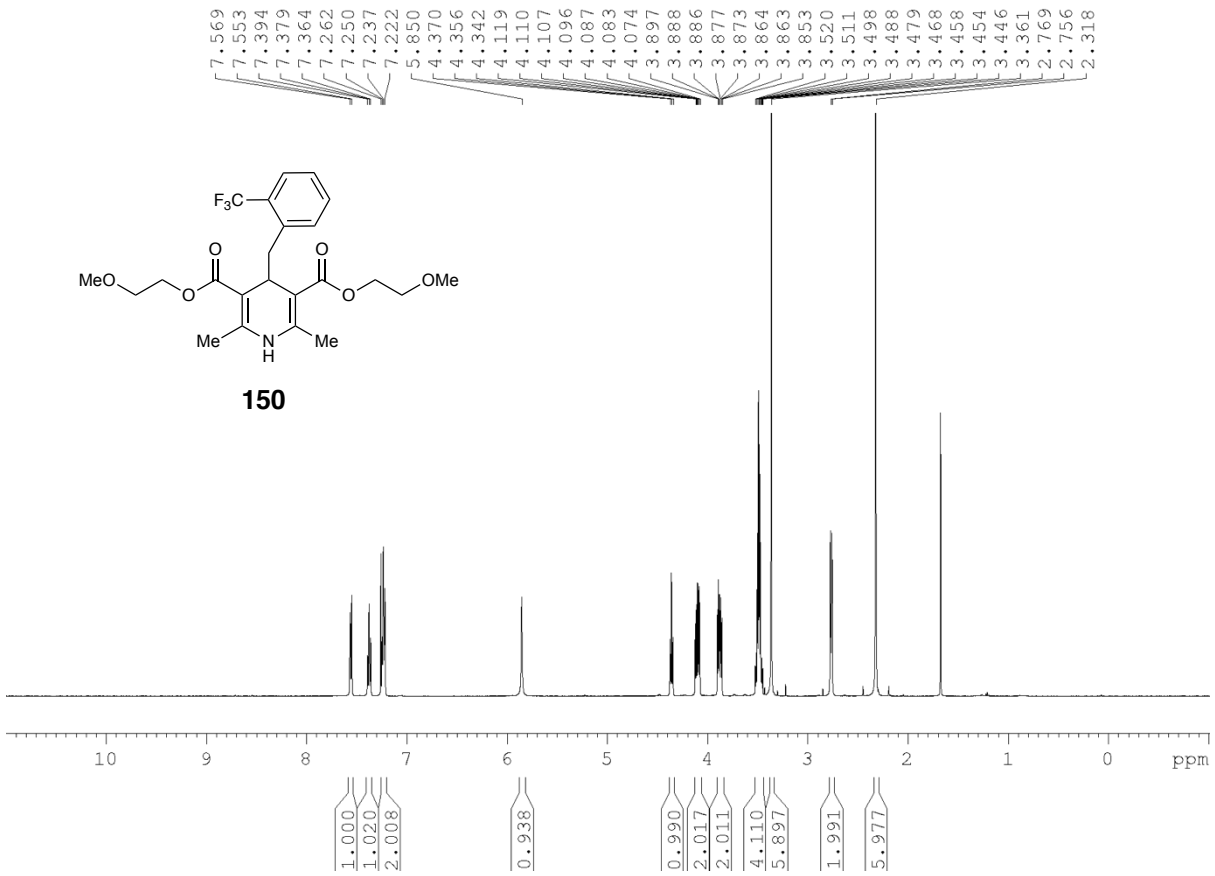



9.764
9.760
9.755
7.627
7.606
7.333
7.315
7.296
7.259
7.247
7.230
7.204
7.186
7.166
3.867
3.863
Br
O
143
10
9
8
7
6
5
4
3
2
1
0
ppm
1.000
1.047
1.129
1.220
1.093
2.267

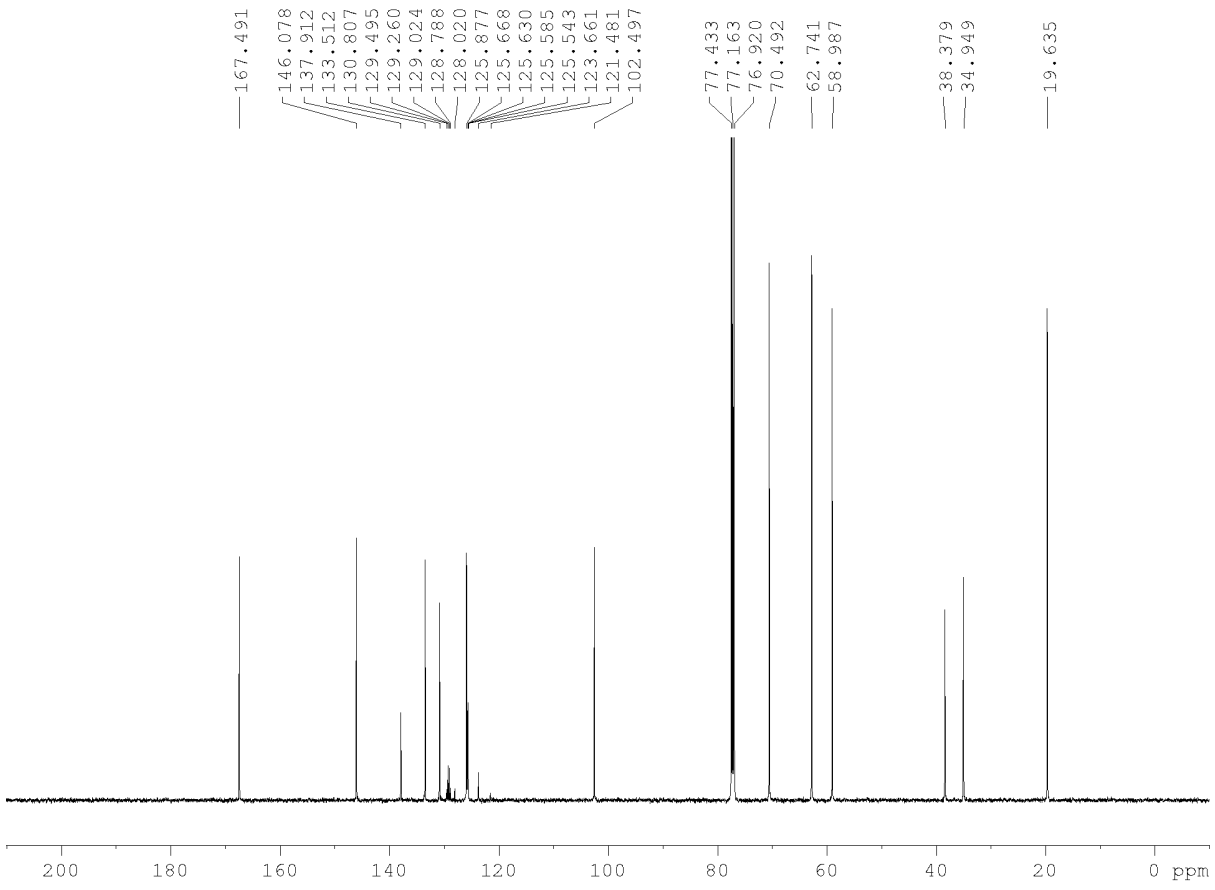

198.401
133.161
132.777
131.899
129.412
128.022
125.113
77.478
77.158
76.842
50.641
200
180
160
140
120
100
80
60
40
20
0 ppm

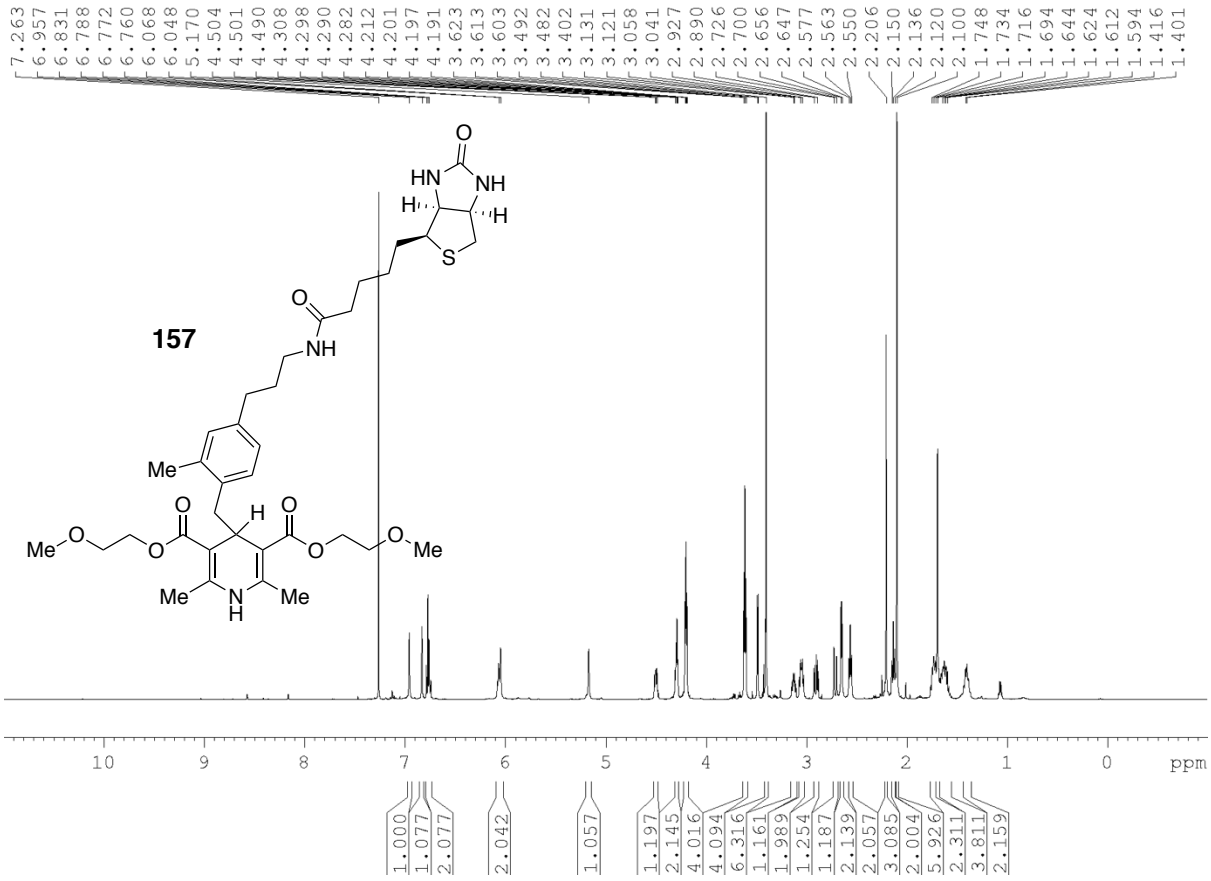

7.455
7.439
7.262
7.139
7.137
7.124
7.122
7.109
7.107
7.049
7.046
7.034
7.031
7.010
7.007
6.994
6.991
6.979
6.976
5.556
4.389
4.376
4.364
4.187
4.177
4.167
4.164
4.153
4.143
4.040
4.030
4.021
4.017
4.007
3.997
3.578
3.568
3.557
3.388
2.807
2.794
2.241
Br
O
O
MeO
O
O
OMe
Me
N
H
Me
149
10
9
8
7
6
5
4
3
2
1
0
ppm
1.000
1.027
1.022
1.018
0.974
1.002
2.057
2.035
4.074
5.950
2.010
6.044

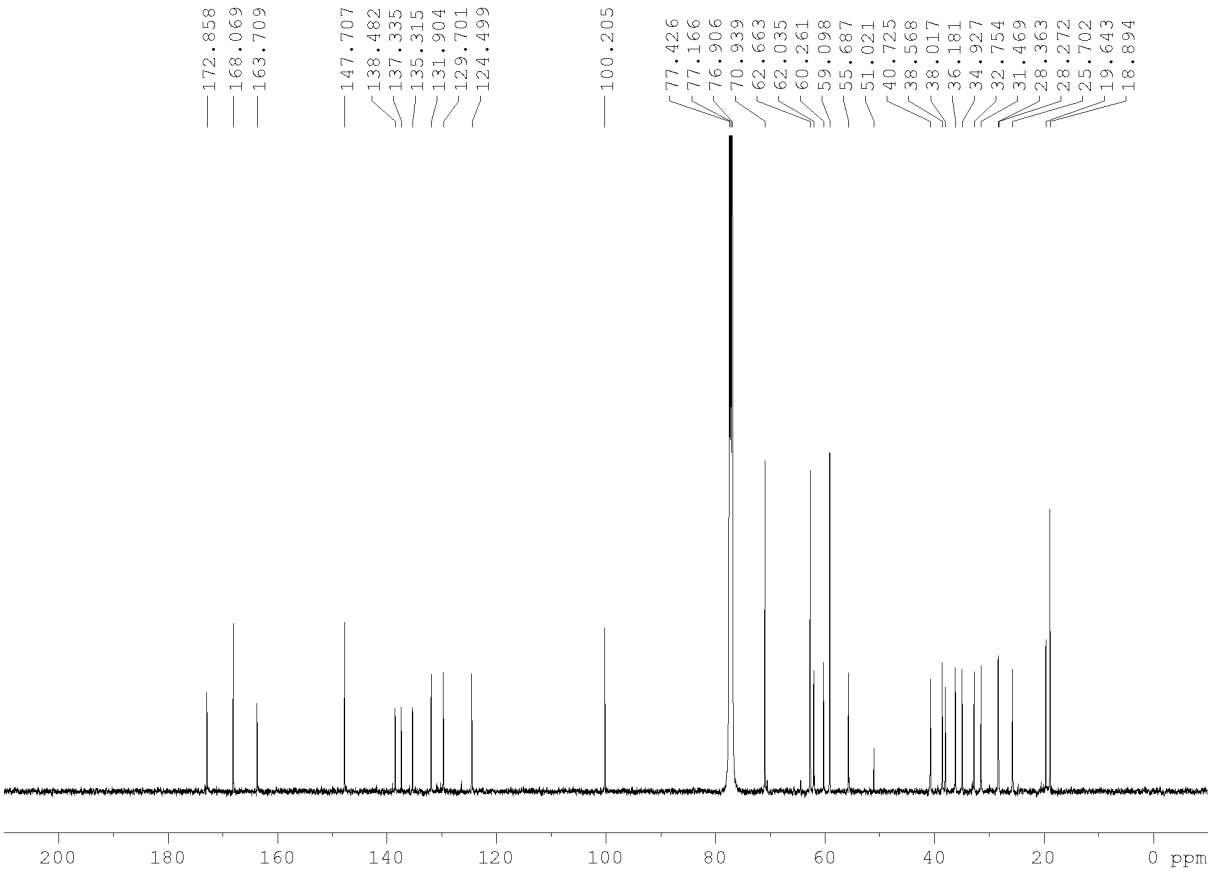

167.566
146.198
138.806
132.687
132.295
127.547
126.421
125.760
101.798
77.443
77.174
76.929
70.761
62.856
59.086
41.194
34.292
19.606
200
180
160
140
120
100
80
60
40
20
0 ppm

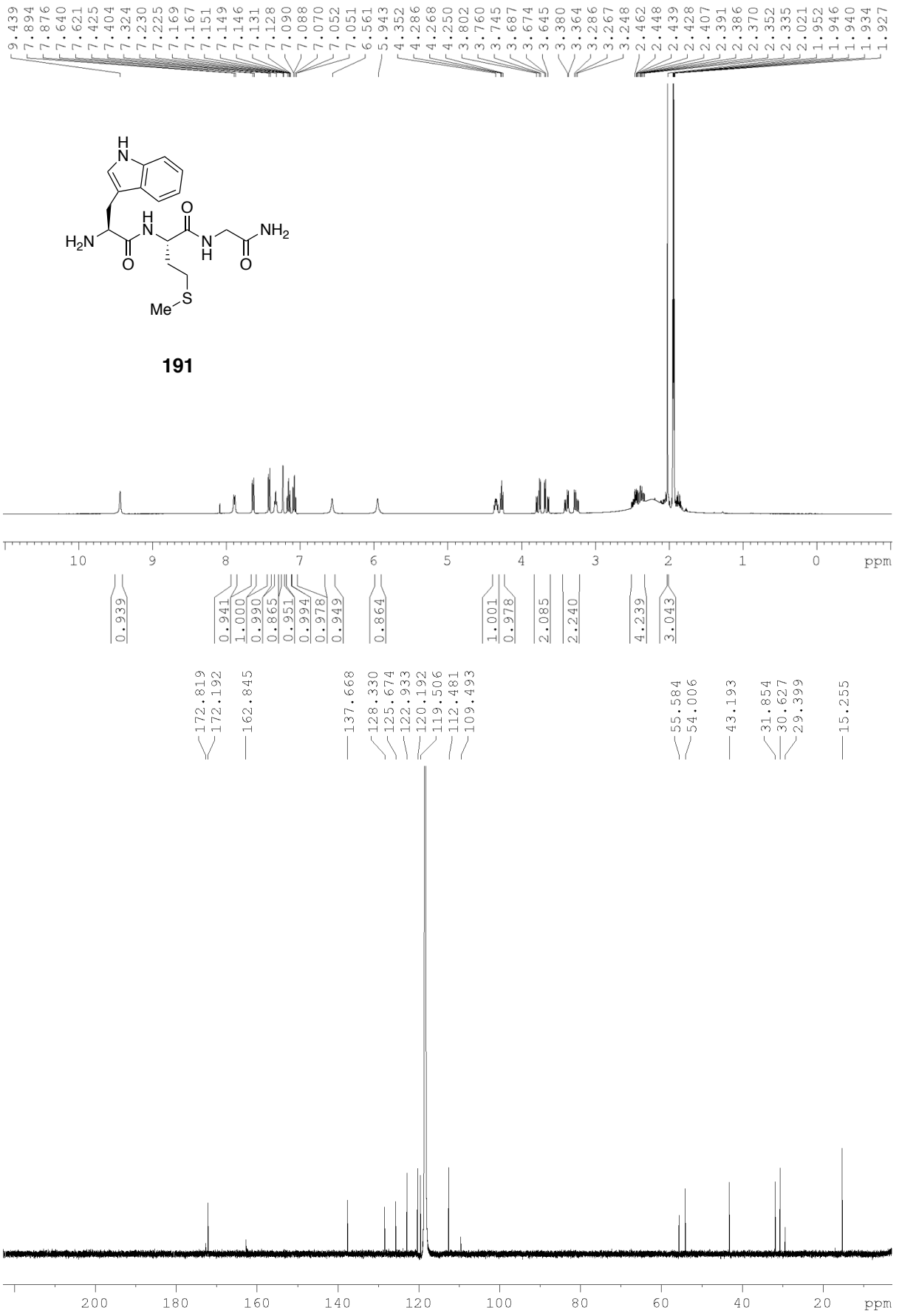

7.569
7.553
7.394
7.379
7.364
7.262
7.250
7.237
7.222
5.850
4.370
4.356
4.342
4.119
4.110
4.107
4.096
4.087
4.083
4.074
3.897
3.888
3.886
3.877
3.873
3.864
3.863
3.853
3.520
3.511
3.498
3.488
3.479
3.468
3.458
3.454
3.446
3.361
2.769
2.756
2.318
F3C
O
O
MeO
O
O
OMe
Me
N
H
Me
10
9
8
7
6
5
4
3
2
1
0
ppm
1.000
1.020
2.008
0.938
0.990
2.017
2.011
4.110
5.897
1.991
5.977

150


167.491
146.078
137.912
133.512
130.807
129.495
129.260
129.024
128.788
128.020
125.877
125.668
125.630
125.585
125.543
123.661
121.481
102.497
77.433
77.163
76.920
70.492
62.741
58.987
38.379
34.949
19.635
200
180
160
140
120
100
80
60
40
20
0 ppm


7.263
6.957
6.831
6.788
6.772
6.760
6.068
6.048
5.170
4.504
4.501
4.490
4.308
4.298
4.290
4.282
4.212
4.201
4.197
4.191
3.623
3.613
3.603
3.492
3.482
3.402
3.131
3.121
3.058
3.041
2.927
2.890
2.726
2.700
2.656
2.647
2.577
2.563
2.550
2.206
2.150
2.136
2.120
2.100
1.748
1.734
1.716
1.694
1.644
1.624
1.612
1.594
1.416
1.401
157
HN
NH
H
H
S
O
O
NH
Me
O
H
O
Me
O
O
O
Me
Me
N
H
Me
10
9
8
7
6
5
4
3
2
1
0
ppm
1.000
1.077
2.077
2.042
1.057
1.197
2.145
4.016
4.094
6.316
1.161
1.989
1.254
1.187
2.139
2.057
3.085
2.004
5.926
2.311
3.811
2.159


172.858
168.069
163.709
147.707
138.482
137.335
135.315
131.904
129.701
124.499
100.205
77.426
77.166
76.906
70.939
62.663
62.035
60.261
59.098
55.687
51.021
40.725
38.568
38.017
36.181
34.927
32.754
31.469
28.363
28.272
25.702
19.643
18.894
200
180
160
140
120
100
80
60
40
20
0 ppm


9.439
7.894
7.876
7.640
7.621
7.425
7.404
7.324
7.230
7.225
7.169
7.167
7.151
7.149
7.146
7.131
7.128
7.090
7.088
7.070
7.052
7.051
6.561
5.943
4.352
4.286
4.268
4.250
3.802
3.760
3.745
3.687
3.674
3.645
3.380
3.364
3.286
3.267
3.248
2.462
2.448
2.439
2.428
2.407
2.391
2.386
2.370
2.352
2.335
2.021
1.952
1.946
1.940
1.934
1.927
H
N
H2N
O
H
N
O
N
H
O
NH2
S
Me
10
9
8
7
6
5
4
3
2
1
0
ppm
0.939
0.941
1.000
0.990
0.865
0.951
0.994
0.978
0.949
0.864
1.001
0.978
2.085
2.240
4.239
3.043
172.819
172.192
162.845
137.668
128.330
125.674
122.933
120.192
119.506
112.481
109.493
55.584
54.006
43.193
31.854
30.627
29.399
15.255
200
180
160
140
120
100
80
60
40
20
ppm

191

Appendix II

Published work

LETTER

https://doi.org/10.1038/s41586-018-0608-y

A protein functionalization platform based on selective reactions at methionine residues

Michael T. Taylor[1], Jennifer E. Nelson[1], Marcos G. Suero[1] & Matthew J. Gaunt[1]*

Nature has a remarkable ability to carry out site-selective post-translational modification of proteins, therefore enabling a marked increase in their functional diversity[1]. Inspired by this, chemical tools have been developed for the synthetic manipulation of protein structure and function, and have become essential to the continued advancement of chemical biology, molecular biology and medicine. However, the number of chemical transformations that are suitable for effective protein functionalization is limited, because the stringent demands inherent to biological systems preclude the applicability of many potential processes[2]. These chemical transformations often need to be selective at a single site on a protein, proceed with very fast reaction rates, operate under biologically ambient conditions and should provide homogeneous products with near-perfect conversion[2–7]. Although many bioconjugation methods exist at cysteine, lysine and tyrosine, a method targeting a less-explored amino acid would considerably expand the protein functionalization toolbox. Here we report the development of a multifaceted approach to protein functionalization based on chemoselective labelling at methionine residues. By exploiting the electrophilic reactivity of a bespoke hypervalent iodine reagent, the *S*-Me group in the side chain of methionine can be targeted. The bioconjugation reaction is fast, selective, operates at low-micromolar concentrations and is complementary to existing bioconjugation strategies. Moreover, it produces a protein conjugate that is itself a high-energy intermediate with reactive properties and can serve as a platform for the development of secondary, visible-light-mediated bioorthogonal protein functionalization processes. The merger of these approaches provides a versatile platform for the development of distinct transformations that deliver information-rich protein conjugates directly from the native biomacromolecules.

The sheer structural diversity of the proteome in any single organism means that no one protein functionalization method is likely to provide a universal solution for the preparation of protein constructs[8–11] (Fig. 1a). Although encoded by the AUG start codon at the beginning of protein synthesis, methionine is often post-translationally excised and thus has a low abundance in proteins (around 2%). It is also frequently used as a replacement for hydrocarbon-containing residues[12]. The limited function of methionine, being mainly responsible for protection against oxidative stress, compared to other residues means that its functionalization is less likely to impair protein function[13]. Targeting methionine would not only provide a distinct bioconjugation approach but also, using our strategy, the resulting methionine conjugates would provide exploitable intrinsic reactivity such that they could lead to the rapid synthesis of diverse, functional protein constructs from native proteins.

So far, there has been only one effective method reported for bioconjugation at methionine, in which the oxidation of thioethers with oxaziridine reagents provided the basis for an elegant bio-inspired strategy to form stable protein-bound sulfoximines[14]. We aimed to target the polarizable thioether on the methionine side chain with a suitable electrophile to form a cationic sulfonium species, selectively installing a versatile payload and distinct functionality at a methionine residue, thereby providing a fundamentally different bioconjugation approach[8–11] (Fig. 1a). Methionine residues react with cyanogen bromide (CNBr); however, such a process cannot function as a bioconjugation method because the instability of the resulting cyano-substituted sulfonium cation triggers cleavage of the protein backbone[15]. Conversely, methionine residues undergo slow *S*-alkylation reactions with iodoacetamide or with other benzyl-derived electrophiles to form relatively stable trialkyl sulfonium cations[16–19]. The need to strike a balance between the stability of the protein-bound sulfonium cations, the compatibility of the reaction conditions, and the reaction rate of the thioether with suitable electrophiles has, so far, precluded the development of an effective alkylation-based method for bioconjugation at methionine. Guided by these limitations, we posited that a distinct class of electrophile, based on the hypervalent iodine scaffold of λ^3-iodanes[20], could make for a functioning bioconjugation process at methionine. Tailoring the substituents and counteranion on the I(III) atom should enable us to tune both the reactivity of the polarizable I(III) nodal centre, to dovetail with the electron lone pair of the thioether, and also the stability of the resulting sulfonium conjugate, through modulation of the electronic features of the groups directly attached to the cationic sulfur motif. We noted that a structurally remarkable iodonium salt (**1** in Fig. 1, R = Et and X = OTf (trifluoromethanesulfonate, also known as triflate)) reacts rapidly with dimethylsulfide (the simplest possible mimic of the thioether motif in methionine) to form a sulfonium adduct[21,22]. Successful reaction of this iodonium salt with methionine would not only represent a distinct method for bioconjugation, but also deliver a high-energy conjugate equipped with reactive 'on-protein' groups that could serve as a basis for designing new transformations towards protein constructs with diverse functionality (Fig. 1b).

We first examined the reaction between dipeptide **2a** and iodonium triflate **1a** (Fig. 2a). We observed the formation of the desired sulfonium conjugate **3a**, although it was clear from the low yield (27%) that the iodonium salt was poorly stable in aqueous solution. By tailoring the aryl group of the iodonium salt (to the electron-deficient 2,4-difluorobenzene) and replacing the triflate counteranion with tetrafluoroborate, we found that a readily prepared reagent **1b** displayed superior physical properties (half-life in water is >50 h). Treatment of reagent **1b** with dipeptide **2a** gave 72% of the desired product **3b** accompanied by the corresponding sulfoxide (not shown) after reaction for 30 min.

Moving to a more complex substrate, the GLP-1 receptor agonist exenatide (Byetta, **2b**, a 39-residue helical polypeptide containing a single mid-chain methionine, Fig. 2b), we found that treatment of a 100 μM aqueous solution with **1b** led to decomposition of the polypeptide. A key breakthrough revealed that the addition of a low concentration of thiourea (20 mM) resulted in a substantial improvement of the reaction, such that sulfonium conjugate **4a** was formed with 68% conversion in less than 2 min, accompanied by non-specific oxidation and labelling. Further improvements could be made by performing the reaction in the presence of TEMPO ((2,2,6,6-tetramethylpiperidin-1-yl)oxyl, 10 mM), which minimized

[1]Department of Chemistry, University of Cambridge, Cambridge, UK. *e-mail: mjg32@cam.ac.uk

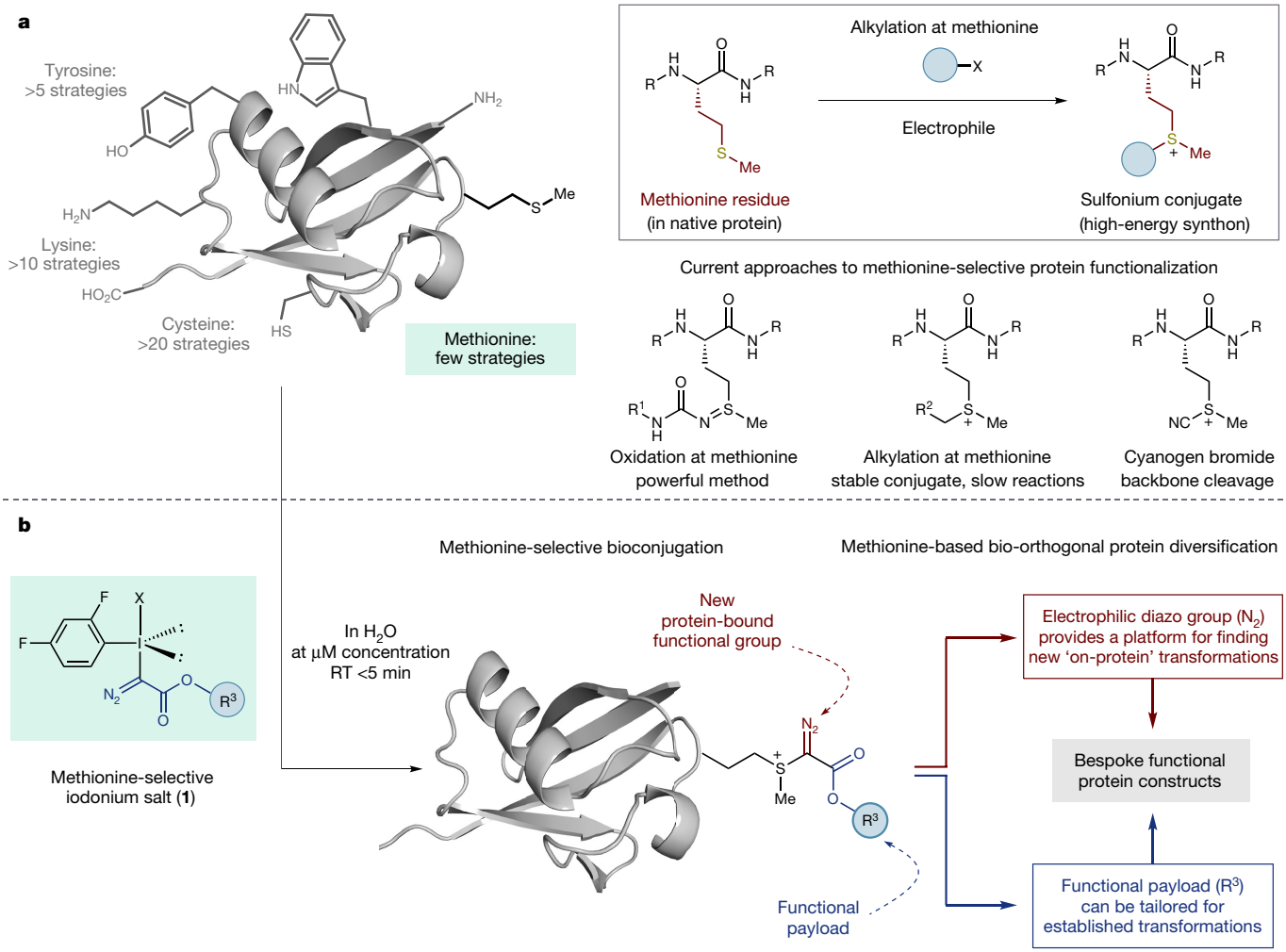


Fig. 1 | The development of a methionine-selective protein functionalization strategy. a, Existing protein functionalization strategy, and the potential for methionine-selective bioconjugation. R = peptide or protein; R^1 = various organic groups; R^2 = aryl or ester group. **b**, Functionalized hypervalent iodine reagents enable methionine-selective protein modification, leading to methionine-based bioorthogonal protein diversification. X = leaving group; R^3 = functional payload.

the formation of oxidative by-products, and adding aqueous formic acid solution (5 mM, approximately pH 3), which reduced the formation of non-specifically labelled by-products to trace levels. Finally, we found that the labelling process proceeded effectively when conducted in distilled water. Routine analysis of electrospray ionization mass spectrometry (ESI-MS) and tandem mass spectrometry (MS/MS) fragmentation data confirmed selective reaction at methionine and, although the concentration of thiourea is well below the levels needed for protein denaturation, it was confirmed by circular dichroism spectra of exenatide conjugate **4a** that the characteristic helical structure was retained (see Supplementary Fig. 10). Although we are not yet certain of the role of thiourea, it is important to note that its presence appears to be fundamental in providing a bioconjugation process at reaction rates needed for transformation on complex biomolecules. With these refined conditions, the thiourea-accelerated bioconjugation was almost instantaneous at exenatide concentrations ranging from

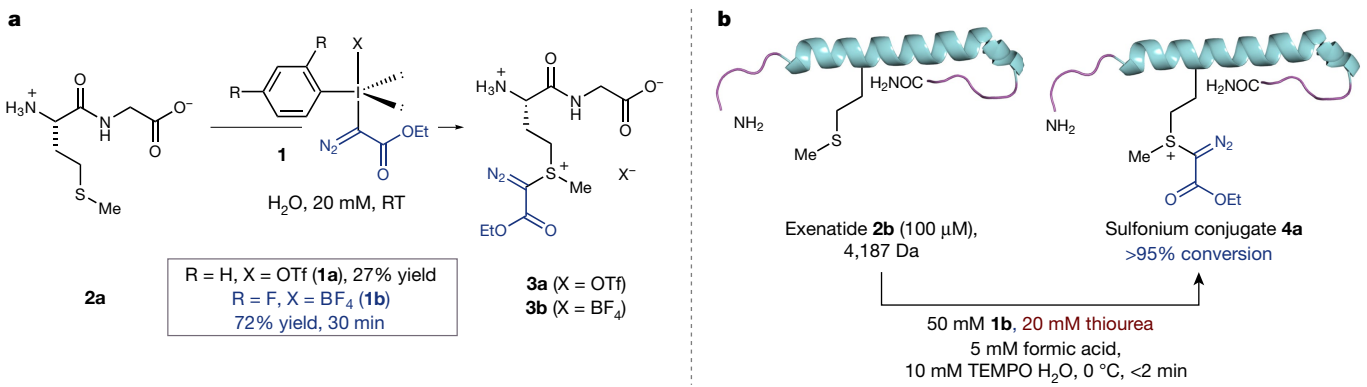


Fig. 2 | Evolution of a methionine-selective bioconjugation strategy. a, Initial results for functionalization at methionine with hypervalent iodine reagents. **b**, Optimal process for the thiourea-accelerated methionine-selective bioconjugation of exenatide.

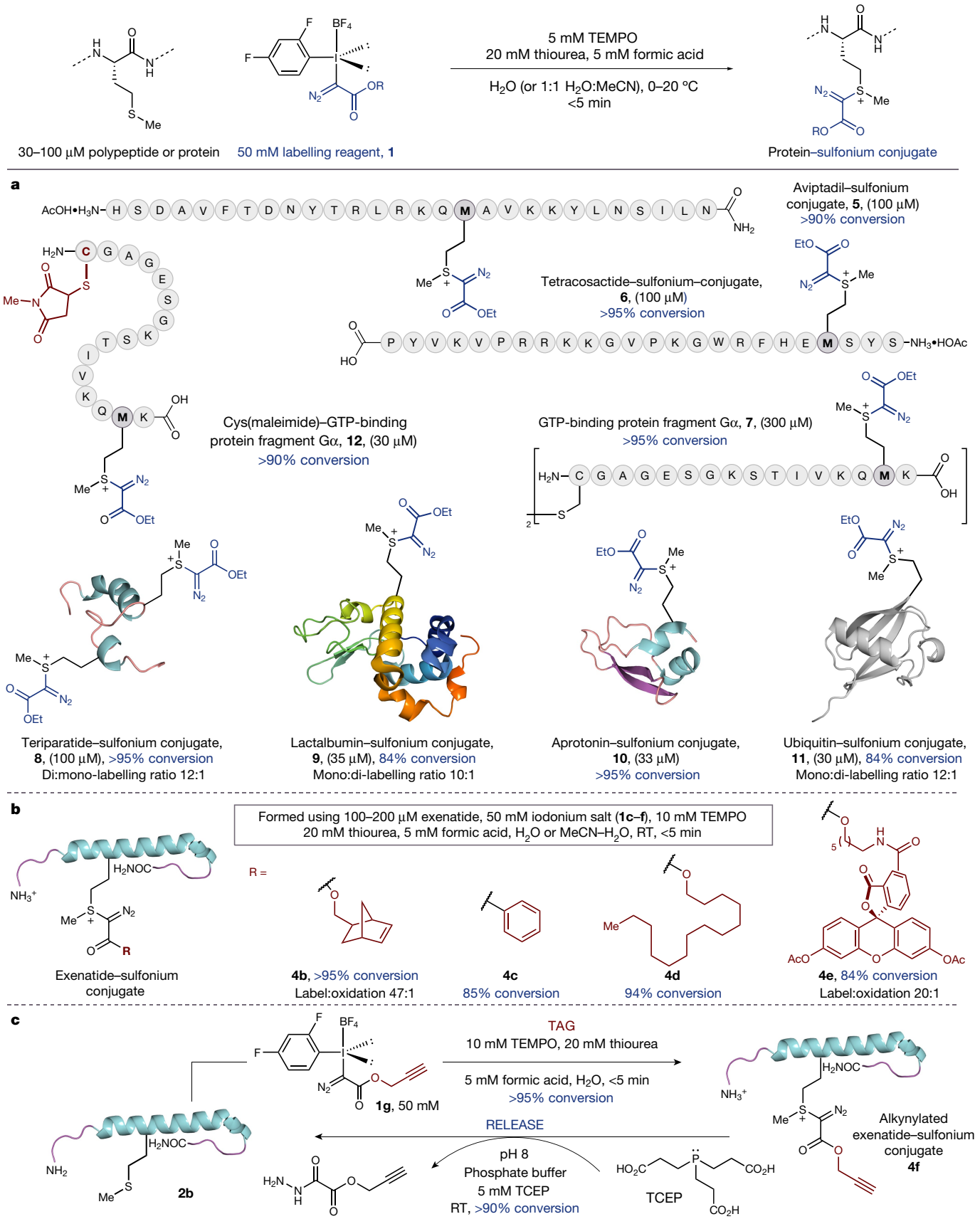


Fig. 3 | Scope of the methionine-selective bioconjugation strategy. a, Range of polypeptide or protein substrates. **b**, Functionalized iodonium reagents compatible with the bioconjugation. **c**, A stimuli-responsive strategy for reversing methionine bioconjugation.

5–500 μM without compromising the conversion, and a larger, milligram-scale reaction enabled the purification of conjugate **4a** by semi-preparative high-performance liquid chromatography to give a 79% yield of isolated product. Given the similarity of the exenatide conjugate to the intermediate that would arise from reaction of the polypeptide with cyanogen bromide (leading to peptide cleavage),

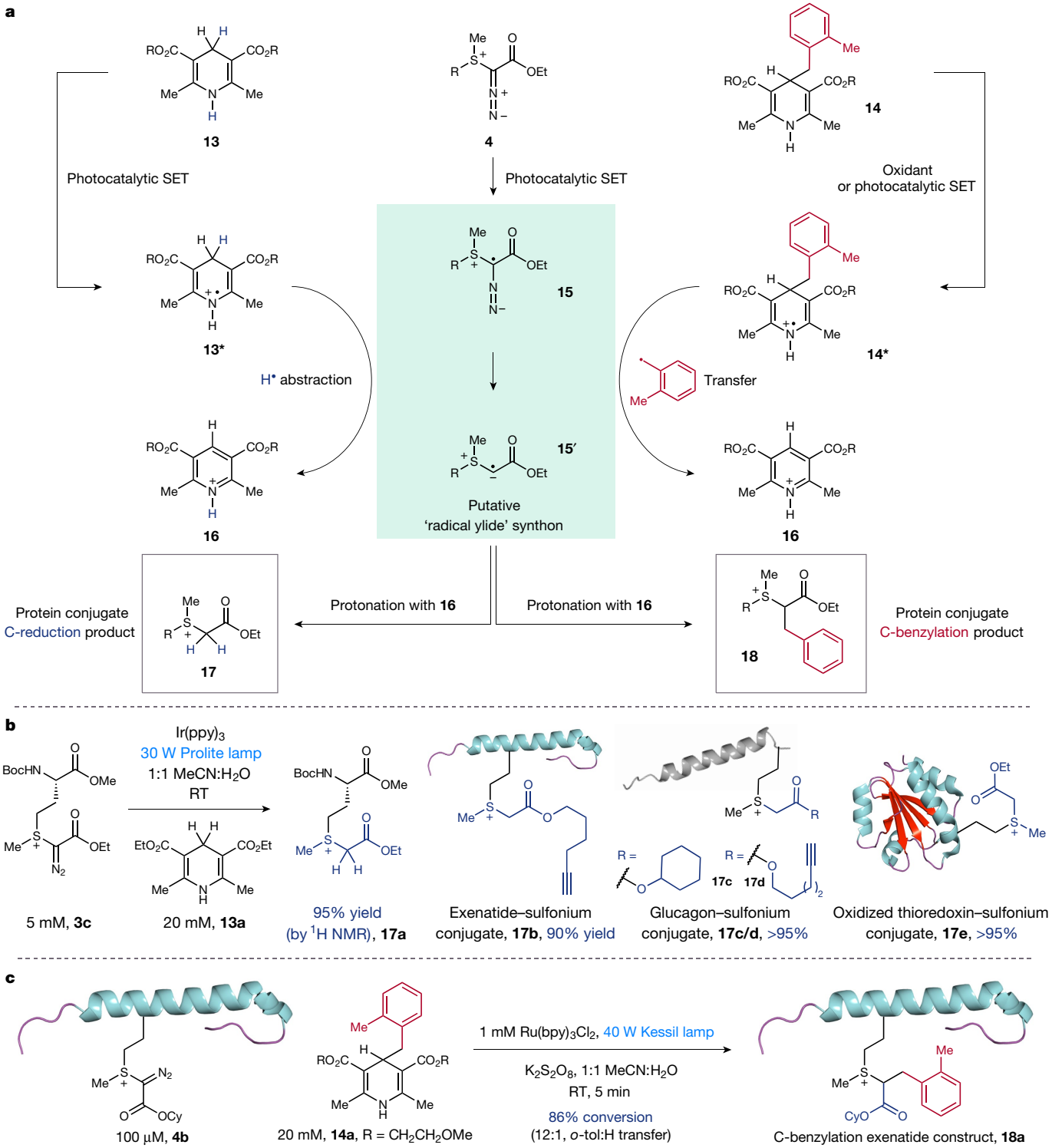


Fig. 4 | Exploiting the multi-faceted reactivity of the protein–sulfonium conjugate. a, A photocatalytic design plan for secondary protein diversification. SET, single-electron transfer. **b**, A system for photoredox-mediated reduction of the sulfonium conjugate and examples of the substrate scope. Notably, the reduction of **3c** proceeds in light, without the iridium photocatalyst, but the yield is greatly reduced. **c**, Secondary protein functionalization via photoredox radical cross-coupling between the diazo sulfonium–protein conjugate and the C-4 substituted Hantzsch ester.

it is remarkable that **4a** has a half-life in water of over 100 h. We believe that the observed stability of **4a** is a result of the ethyl diazoacetate motif imparting a lesser electron-withdrawing effect (than the cyano group) on the sulfonium salt, which in turn leads to a species that is less reactive to attack by the proximal carbonyl group and is, therefore, a more stable conjugate.

The optimized conditions, using **1b**, were used to evaluate the substrate scope (Fig. 3a). Random coil polypeptides that contain methionine, such as aviptadil and tetracosactide, are efficiently converted to the sulfonium conjugates **5** and **6** respectively. GTP-binding protein fragment Gα, which contains a free sulfhydryl group in an N-terminal cysteine residue, underwent smooth methionine labelling to **7** with concomitant oxidative disulfide formation. Teriparatide, a polypeptide containing two methionines, formed bis-sulfonium conjugate (**8**) with good conversions. α-Lactalbumin, a globular protein with a readily oxidized methionine residue,

undergoes bioconjugation to **9** with minimal competitive oxidation. Aprotinin (Trasylol) also forms the corresponding sulfonium conjugate **10** in good conversion to product. Particularly notable is the observation that two cysteine disulfide linkages within the structure of aprotinin (and the labelling to **7**) are not affected by **1b**, with a single methionine-derived conjugate obtained in high conversion. When the methionine residues are buried within the tertiary structure of the protein, for example with RNA-ase B, the bioconjugation does not occur, highlighting the inherent selectivity of the labelling process for exposed rather than inaccessible thioether groups. This feature is highlighted by the case of ubiquitin: the N-terminal methionine residue has only moderate surface exposure and can slow down the rate of functionalization to the point where oxidation becomes competitive. Accordingly, reaction under deoxygenated conditions enabled efficient conversion to the labelled product (84%), with a 10:1 label:oxidation ratio (**11**). The methionine-selective bioconjugation strategy effectively functionalizes a range of polypeptides and proteins in high conversion at micromolar concentrations and in very short reaction times. Given that **1b** is a carbon electrophile with two of the best leaving groups known to organic chemists, it would be thought to be very reactive; it is therefore remarkable that this species selectively engages the moderately nucleophilic methionine residue in the presence of competitively nucleophilic and oxidizable amino acid residues. Bioconjugation strategies that target other amino acids should, therefore, be compatible with, and complementary to, our methionine-functionalization process. To exemplify this, we showed that the GTP-binding protein fragment Gα could be first labelled at cysteine, using a maleimide derivative[23], and then conjugated at methionine using iodonium salt **1b** to form **12**. The methionine-selective process does not interfere with the cysteine–maleimide motif, which itself contains a thioether linkage, thereby highlighting possible applications towards multi-site heterolabelling of proteins[24].

The modular synthesis of the hypervalent iodine reagents enables facile incorporation of different acyl groups attached to the diazo motif, allowing the transfer of a range of functional payloads to proteins (Fig. 3b). Functional groups relevant to other bioorthogonal reactions are readily tolerated in both the reagent synthesis and the methionine labelling, smoothly affording sulfonium conjugates **4b** and **4c** with 95% and 85% conversion to product, respectively. Biochemical reporter groups, such as myristyl- and fluorescein-derived esters **4d** and **4e**, are also readily transferred to exenatide. Notably, we found that sulfonium conjugates of exenatide (such as **4f**) underwent reaction with the tertiary phosphine TCEP (tris(2-carboxyethyl)phosphine), a standard biochemical reagent, resulting in the cleavage of the labelling group and return of the parent exenatide **2b** in >90% conversion[25] (Fig. 3c); the cleavage reaction also works for conjugates **5**, **6** and **8** in comparable conversions and provides a stimulus-responsive means of removing the methionine label.

Next, we turned our attention to exploring the multifaceted reactivity that we anticipated would be intrinsic to the high-energy methionine-derived conjugate. The electrophilicity of the diazo sulfonium conjugate **4** prompted us to investigate the single-electron transfer chemistry of this reactive motif enabled by visible-light photocatalysis[26,27]. The addition of a single electron to the diazo sulfonium conjugate **4** could result in intermediate **15**, which, upon cleavage of the C–N_2 bond, would form a putative radical ylide synthon **15′** (Fig. 4a). We visualized two pathways through which we could exploit the reactivity of the previously unexplored radical ylide. First, combining the radical ylide with Hantzsch ester **13*** (from **13**) may lead to a reduction process resulting in the generation of a trialkylsulfonium motif, which would impart enhanced chemical stability to the protein conjugate. Furthermore, the use of a C-4 benzylated Hantzsch ester derivative (**14**, an established precursor for a benzyl radical)[28] to intercept the radical ylide species would lead to a C-benzylation product that could be used to introduce functional diversity to the protein conjugate. We screened a range of photocatalysts under visible-light conditions. When **3c** was irradiated with a 30 W lamp in the presence of *fac*-Ir(ppy)$_3$ (ppy, 2-phenylpyridinato) and the Hantzsch ester **13**[29,30], we observed the formation of the reduced trialkylsulfonium product **17a** in 95% yield (determined by ^{1}H NMR, Fig. 4b). Using these conditions, we showed that a range of sulfonium–protein conjugates, including exenatide, glucagon and thioredoxin derivatives, are reduced to stable trialkylsulfonium species with excellent conversions (**17b–e**, see Supplementary Information). Notably, reduction of a thioredoxin derivative[31] to its trialkylsulfonium–protein congener **17e** proceeds in high conversion without affecting its labile disulfide linkage, which serves to highlight the mild nature of this protocol. Additionally, the methionine bioconjugation and photoreduction steps can be carried out in a one-pot operation, which considerably simplifies the overall process without compromising the yield or purity of the trialkylsulfonium product.

In testing the viability of the proposed C-benzylation using a modified Hantzsch ester derivative[28], we found that treatment of the exenatide–sulfonium conjugate **4b** with *o*-tolyl Hantzsch ester derivative **14a**, under slightly modified photocatalytic conditions, led to cross-coupling and the formation of the C-ligation product **18a** in high conversion. This unique bioorthogonal protein functionalization reaction not only represents a synthetic radical–radical cross-coupling using a polypeptide scaffold, but also provides a platform for bioorthogonal protein diversification wherein two distinct functionalities could be introduced sequentially at the same amino acid residue.

In summary, through the merger of methionine-selective bioconjugation and a new visible-light-mediated photocatalytic reaction platform, information-rich synthetic constructs can be rapidly assembled by a two-step protocol directly from native proteins. The reactivity inherent to the methionine conjugate distinguishes the bioconjugation process from other methods. Moreover, the capacity for functional diversification, by tailoring the hypervalent iodine reagent and modified Hantzsch-ester derivative, means that highly functional protein conjugates could be made readily available directly from native proteins.

Data availability

The data that support the findings of this study are available within the paper and its Supplementary Information. Raw data are available from the corresponding author on reasonable request.

Received: 27 February 2017; Accepted: 21 August 2018;
Published online 15 October 2018.

1. Walsh, C. T., Garneau-Tsodikova, S. & Gatto, G. J. Jr. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew. Chem. Int. Ed.* **44**, 7342–7372 (2005).
2. Sletten, E. M. & Bertozzi, C. R. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed.* **48**, 6974–6998 (2009).
3. Spicer, C. D. & Davis, B. G. Selective chemical protein modification. *Nat. Commun.* **5**, 4740 (2014).
4. Koniev, O. & Wagner, A. Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation. *Chem. Soc. Rev.* **44**, 5495–5551 (2015).
5. Dawson, P. E. & Kent, S. B. H. Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **69**, 923–960 (2000).
6. Lang, K. & Chin, J. W. Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem. Rev.* **114**, 4764–4806 (2014).
7. Wang, L., Xie, J. & Schultz, P. G. Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 225–249 (2006).
8. Vinogradova, E. V., Zhang, C., Spokoyny, A. M., Pentelute, B. L. & Buchwald, S. L. Organometallic palladium reagents for cysteine bioconjugation. *Nature* **526**, 687–691 (2015).
9. Wright, T. H. et al. Posttranslational mutagenesis: a chemical strategy for exploring protein side-chain diversity. *Science* **354**, aag1465 (2016).
10. Yang, A. et al. A chemical biology route to site-specific authentic protein modifications. *Science* **354**, 623–626 (2016).
11. Abegg, D. et al. Proteome-wide profiling of targets of cysteine reactive small molecules by using ethynyl benziodoxolone reagents. *Angew. Chem. Int. Ed.* **54**, 10852–10857 (2015).
12. Levine, R. L., Moskovitz, J. & Stadtman, E. R. Oxidation of methionine in proteins: roles in antioxidant defense and cellular regulation. *IUBMB Life* **50**, 301–307 (2000).

13. Cowie, D. B., Cohen, G. N., Bolton, E. T. & De Robichon-Szulmajster, H. Amino acid analog incorporation into bacterial proteins. *Biochim. Biophys. Acta* **34**, 39–46 (1959).
14. Lin, S. et al. Redox-based reagents for chemoselective methionine bioconjugation. *Science* **355**, 597–602 (2017).
15. Gross, E. & Witkop, B. Nonenzymatic cleavage of peptide bonds: the methionine residues in bovine pancreatic ribonuclease. *J. Biol. Chem.* **237**, 1856–1860 (1962).
16. Gundlach, H. G., Stein, W. H. & Moore, S. The nature of the amino acid residues involved in the inactivation of ribonuclease by iodoacetate. *J. Biol. Chem.* **234**, 1754–1760 (1959).
17. Vithayathil, P. J. & Richards, F. M. Modification of the methionine residue in the peptide component of ribonuclease-S. *J. Biol. Chem.* **235**, 2343–2351 (1960).
18. Kramer, J. R. & Deming, T. J. Preparation of multifunctional and multireactive polypeptides via methionine alkylation. *Biomacromolecules* **13**, 1719–1723 (2012).
19. Kramer, J. R. & Deming, T. J. Reversible chemoselective tagging and functionalization of methionine containing peptides. *Chem. Commun.* **49**, 5144–5146 (2013).
20. Stang, P. J. & Zhdankin, V. V. Organic polyvalent iodine compounds. *Chem. Rev.* **96**, 1123–1178 (1996).
21. Weiss, R., Seubert, J. & Hampel, F. α-Aryliodonio diazo compounds: S_N reactions at the α-C atom as a novel reaction type for diazo compounds. *Angew. Chem. Int. Edn Engl.* **33**, 1952–1953 (1994).
22. Schnaars, C., Hennum, M. & Bonge-Hansen, T. Nucleophilic halogenations of diazo compounds, a complementary principle for the synthesis of halodiazo compounds: experimental and theoretical studies. *J. Org. Chem.* **78**, 7488–7497 (2013).
23. Kim, Y. et al. Efficient site-specific labeling of proteins via cysteines. *Bioconjug. Chem.* **19**, 786–791 (2008).
24. Mülhberg, M. et al. Orthogonal dual-modification of proteins for the engineering of multivalent protein scaffolds. *Beilstein J. Org. Chem.* **11**, 784–791 (2015).
25. Staudinger, H. & Lüscher, G. Über darstellung und reaktionen von phosphazinen. *Helv. Chim. Acta* **5**, 75–86 (1922).
26. Prier, C. K., Rankic, D. A. & MacMillan, D. W. C. Visible light photoredox catalysis with transition metal complexes: applications in organic synthesis. *Chem. Rev.* **113**, 5322–5363 (2013).
27. Chen, Y., Kamlet, A. S., Steinman, J. B. & Liu, D. R. A biomolecule-compatible visible-light-induced azide reduction from a DNA-encoded reaction-discovery system. *Nat. Chem.* **3**, 146–153 (2011).
28. Huang, W. & Cheng, X. Hantzsch esters as multifunctional reagents in visible-light photoredox catalysis. *Synlett* **28**, 148–158 (2017).
29. Fukuzumi, S., Hironaka, K. & Tanaka, T. Photoreduction of alkyl halides by an NADH model compound. An electron-transfer chain mechanism. *J. Am. Chem. Soc.* **105**, 4722–4727 (1983).
30. Hedstrand, D. M., Kruizinga, W. H. & Kellog, R. M. Light induced and dye accelerated reductions of phenacyl onium salts by 1,4-dihydropyridines. *Tetrahedron Lett.* **19**, 1255–1258 (1978).
31. Krause, G., Lundström, J., Barea, J. L., Pueyo de la Cuesta, C. & Holmgren, A. Mimicking the active site of protein disulfide-isomerase by substitution of proline 34 in *Escherichia coli* thioredoxin. *J. Biol. Chem.* **266**, 9494–9500 (1991).

Acknowledgements We thank M. Nappi and C. Guerot for advice and useful discussions. We thank the Marie Curie Actions program (M.T.T. and M.G.S.), AstraZeneca and EPRSC (J.E.N.), and the European Research Council (ERC-SRG-259711), EPSRC (EP/100548X/1) and the Royal Society (Wolfson Merit Award) for fellowships (M.J.G.). We are grateful to J. Chin, N. Huguen, M. Skehel, H. Lewis and M. Edgeworth for assistance with protein purification and mass spectrometry experiments.

Reviewer information *Nature* thanks A. Spokoyny and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions M.J.G., M.T.T., J.E.N. and M.G.S. conceived the project and designed the experiments. M.J.G., M.T.T., J.E.N. and M.G.S. performed and analysed the experiments. M.J.G., M.T.T. and J.E.N. wrote the paper.

Competing interests The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0608-y.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to M.J.G.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.